Energy Research and Development Division
FINAL PROJECT REPORT

THE DESIGN AND EVALUATION OF A MULTI-TIME SCALE COMPOSITE SOLAR FORECAST SYSTEM

Prepared for: California Energy Commission
Prepared by: AWS Truepower, LLC

PREPARED BY:

Primary Author(s):
    Philippe Beaucage
    John Zack
    Edward Natenberg
    Adam Kankiewicz
    Jan Kleissl
    Seyman Iman Gohari
    Bryan Urquhart
    Richard Perez

    AWS Truepower, LLC
    463 New Karner Road
    Albany, NY 12205

Contract Number: 500-10-057

Prepared for:

California Energy Commission

Michael Sokol
Contract Manager

Aleecia Gutierrez
Office Manager
Energy Generation Research Office

Laurie ten Hope
Deputy Director
ENERGY RESEARCH AND DEVELOPMENT DIVISION

Robert P. Oglesby
Executive Director

ACKNOWLEDGEMENTS

PREFACE

The California Energy Commission Energy Research and Development Division supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The Energy Research and Development Division conducts public interest research, development, and demonstration (RD&D) projects to benefit California.

The Energy Research and Development Division strives to conduct the most promising public interest energy research by partnering with RD&D entities, including individuals, businesses, utilities, and public or private research institutions.

Energy Research and Development Division funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency

- Energy Innovations Small Grants

- Energy-Related Environmental Research

- Energy Systems Integration

- Environmentally Preferred Advanced Generation

- Industrial/Agricultural/Water End-Use Energy Efficiency

- Renewable Energy Technologies

- Transportation

The Design and Evaluation of a Multi-Time Scale Composite Solar Forecast System is the final report for the Research Needs for Utility-Scale Renewable Energy project (contract number 500-10-057 conducted by AWS Truepower. The information from this project contributes to Energy Research and Development Division's Renewable Energy Technologies Program.

For more information about the Energy Research and Development Division, please visit the Energy Commission's website at www.energy.ca.gov/research/ or contact the Energy Commission at 916-327-1551.

ABSTRACT

Minimizing the cost of integrating solar power into the electric grid requires accurate foreknowledge of solar-based electrical production on a wide range of look-ahead time scales. These forecasts allow grid operators to cost effectively manage the impact of its variability. The goal of this project was to assemble and evaluate a composite solar forecast system that optimally integrated four different forecasting methods - sky imagers, satellite and two variants of numerical weather prediction (NWP) models - to achieve the best forecast performance for each look-ahead period. The look-ahead period ranged from 10 minutes to 48 hours. A 48-megawatt (MW) utility-scale photovoltaic (PV) power plant in Nevada was the forecast target. This facility was chosen as the forecast target and evaluation site because its power is fed into the California Independent System Operator (CAISO) grid.

Solar hindcasts, or forecasts generated in hindsight, were produced from each individual method and the performance of each forecast for the target look-ahead period was objectively evaluated using standard metrics. For the intra-hour period (10-30 minutes), the persistence forecasts achieved the lowest error. The satellite-based forecast outperformed the persistence forecasts from 30 minutes up to a 3 hour look-ahead period. The rapid update numerical weather prediction forecast outperformed the satellite method for the 3.5 to 5 hour look-ahead. For the day-ahead period (6-48 hours), most numerical weather prediction-based forecasts were more accurate than the climatological forecasts.

A short-term (10 minutes to 5 hours) and a longer-term (6 to 48 hours) composite forecast method was designed and tested. Both composite forecasts had lower error than any individual methods on average. Therefore, by combining the best available forecast information for each look-ahead time, the composite forecast provides better accuracy and foreknowledge of solar-based electricity production. The cost-effective management of solar power generation requires high forecast accuracy and tools that can capture the variations in energy production at multiple time scales that are critical to situational awareness and grid operational planning (economic dispatch, curtailment, etc.). Significant benefits are also derived from advancements in solar forecasting such as lower electricity cost associated with a higher level of solar penetration and reliability in addition to the reduction in greenhouse gas emissions.

TABLE OF CONTENTS

LIST OF FIGURES

viii

LIST OF TABLES

EXECUTIVE SUMMARY

Introduction

Solar power is a very dynamic energy source and its generation is directly linked to the availability of solar radiation. Solar radiation hitting a given point on the Earth exhibits substantial variability on time scales of days, hours and even minutes, as a result of changes in atmospheric conditions. This variability presents a significant challenge to grid management. Clouds can significantly reduce the power generation from a utility-scale solar plant by shading all or large portions of its array for minutes, hours or days. In some cases, the transition from clear to cloudy conditions can result in a steep, rapid decline in power available to meet the current load on the grid, referred to as a "ramp" event.

Grid operators are confronted with integrating this highly variable, non-dispatchable energy source, especially where solar power is available in high concentration, such as in California. Solar energy production forecasts can be used to provide grid operators with situational awareness, allowing them to maintain grid stability while incorporating high penetrations of solar power. Timely and accurate foreknowledge of minutes-ahead, hour-ahead and day-ahead solar-based generation is critical. As penetration levels increase across California, advanced forecasting methods are anticipated to be one of the most cost-effective tools for the management of non-dispatchable variable power generation on the grid.

Advanced solar forecasting methods incorporate various technologies and data sources to predict energy production, including ground-based, sun-tracking cameras or sky imagers; visible and infrared satellite imagery; and computer-based numerical weather prediction (NWP) models. Solar forecast methods targeted for specific look-ahead periods (e.g. intra-hour, hour-ahead, day-ahead) have been developed and are in operational use by plant operators, balancing authorities and energy schedulers.

Sky imagers, satellite and NWP model approaches each have strengths and weaknesses when applied as stand-alone forecasting applications. Previous research and operational experience indicates that there is no single forecast method that is optimal for all relevant look-ahead periods. An approach that provides the highest possible accuracy over a broad range of time scales by utilizing different forecasting methods has not been implemented or evaluated prior to this research project.

Project Purpose

The purpose of this project was to configure, demonstrate and validate a composite forecast tool that optimally integrates several state-of-the-art solar forecasting methods for California into a singular output to achieve the best forecast performance within the targeted look-ahead periods, ranging from minutes-ahead to day-ahead. Advancements in forecasting techniques for variable power output will facilitate grid integration of utility-scale solar energy at high penetration levels.

Project Approach

A multiple time scale composite forecast system was developed from a set of refined pre-existing forecast methods and evaluated using a year of performance data gathered from the target site. Copper Mountain Solar 1, a utility-scale photovoltaic power plant in Nevada was chosen as the target evaluation site because power generated at this plant is fed into the California Independent System Operator (CAISO) grid. The composite solar forecasting tool was developed from four individual forecast methods: sky imager, satellite, and two different types of NWP. The individual solar forecast methods and the resulting composite tool were evaluated at multiple time scales relevant to grid operations and reserve planning using three different look-ahead periods:

- Intra-hour (10 minutes to 30 minutes)

- Hour-ahead (30 minutes to 5 hours)

- Day-ahead (6-48 hours)

Sky imagers were deployed at the target site to detect cloud movement and sky cover conditions for the purpose of forecasting intra-hour (10-30 minutes) solar radiation and power generation. Satellite imagery provided solar forecasts for the look-ahead period of 10 minutes to 5 hours ahead (intra-hour and hour-ahead) using a cloud motion vector scheme. Forecasts from 30 minutes out to several days ahead were obtained from NWP models that included both the hour-ahead and day-ahead target periods. The composite forecast system was designed to combine each of these methods and assign a weight to each according to its performance for each targeted look-ahead period.

The project was divided into three primary efforts undertaken by research team members, University of California- San Diego (UCSD), Clean Power Research (CPR), MESO, Inc. and AWS Truepower, LLC. Through this project:

- Significant amounts of data were collected, quality-controlled and stored for individual use by each team member. This included plant power generation and on-site irradiance, image data from on-site sky imagers, as well as satellite imagery and the output from numerous NWP models. Regional data sets from airports or other meteorological collection stations were also utilized. Each project team member collected, quality-controlled and maintained their individual datasets for training and validation efforts.

- Forecasts produced using sky imager, satellite, and NWP models were individually evaluated for their target look-ahead periods using data from the 2012 calendar year. The performance of each individual solar forecast method for the target look-ahead periods was objectively evaluated using on-site plant data and standard error metrics.

- The composite forecast system was designed to produce a deterministic and probabilistic forecast by statistically weighting each individual method for each look-ahead period. The relative performance of each individual method over a representative historical sample determined the weighting. That is, a heavier weighting was given to

the method that performed best for each look-ahead period. The composite forecast results for 2012 were compared with the results of each individual method to evaluate the relative differences in accuracy and performance.

Results and Conclusions

The individual performance of each forecast method for the look-ahead period of interest was evaluated relative to each other and standard reference forecasts using persistence and climatology. A persistence forecast assumes that the current weather condition will persist and that future conditions will be the same as the present. A climatological forecast is based solely upon the climatological statistics of a specific site or region. This method involves averaging the solar irradiance or power generation observed at the same time of day (e.g. hour) and time of year (e.g. day and month) over many years to generate the forecast. Even though persistence forecasts cannot predict ramping events, they typically perform well for intra-hour forecast horizons because the magnitude of intra-hour changes is small for a large fraction of the time. Climatological forecasts provide a useful baseline for the day-ahead forecast horizon, i.e. 6-48 hour look-ahead period.

A wide variety of metrics are used for evaluating forecast performance. For this project, the team focused on the two most common metrics: mean bias error and root mean square error (RMSE). The mean bias error indicates whether the forecasting methods are properly calibrated whereas the RMSE is quite sensitive to outliers and therefore large errors in forecasts. Researchers also selected a mix of clear sky, mostly sunny, mostly cloudy and overcast conditions for the initial validation of each forecasting method in the 10 to 30 minute look-ahead period.

For the intra-hour period (10-30 minutes), the persistence forecasts achieved the lowest RMSE when compared to the sky imager and satellite results. However, the RMSE of the persistence forecasts increases with lead time and becomes almost as large as the satellite forecast at a 30-minute lead time.

Under clear sky or overcast conditions, the sky imager method was not able to infer the motion of the clouds based on two consecutive images. The satellite images cover a large area (several hundred kilometers), which means there are typically some cloud structures that can be projected forward in time and may explain why the error of the satellite method remains relatively stable over the 10 to 30 minute lead time. Although the persistence forecast exhibits the best RMSE performance in the 10-30 minute look-ahead period, it will not predict ramping events caused by the boundaries of cloud shadows and therefore its utility is limited when weather conditions are rapidly evolving. These are typically the times that the grid system operators have the greatest need for accurate solar forecasts.

The intra-hour forecast validation was based on a relatively small sample size with a mix of cloud cover conditions that are not exactly representative of the true distribution at the Copper Mountain plant, which may have introduced some bias into the forecast performance evaluation. Additionally, the imagery available from the sky imager unit was limited by the shadowband, which holds the camera above the mirror, blocking its field of view. Poor data

recovery due to soiling as well as manual processes limited the sky imager forecast data available for the composite forecast. Sky imager forecasting is still in development and advancements in technology will likely remedy some of these issues while increasing its potential value to detect very short-term ramp events.

Results for the hour-ahead and day-ahead periods were generated with the satellite method and NWP models, and they were evaluated with respect to irradiance. The NWP forecasts were updated every 2 hours at specific times, unlike the satellite forecasts which were updated every 30 minutes. Therefore, the researchers evaluated the forecasts over concurrent delivery and horizon times, and only considered the 2-hour update cycles as they were the lowest common denominator among both forecasting approaches.

The error in irradiance forecasts varied substantially by time of day and seasonally. At the target plant, all forecast approaches perform well in May and June when clear sky conditions prevail. In July and August, errors are much larger than for other months, likely due to highly variable cloud conditions prevailing during the North American monsoon season, which is more challenging to forecast than prevailing clear sky conditions.

The error growth by look-ahead time indicates that the persistence forecast loses value over the climatological forecast after a 2-hour look-ahead time. For the hour-ahead period, the satellite-based forecast outperformed every forecasting method from 30 minutes up to 3 to 4 hours ahead. The NWP forecasts have higher error values than the satellite forecasts at a 30-minute lead time, but they outperformed the satellite forecasts for the 3.5 to 5 hour look-ahead time. Throughout the 10 minutes to 5 hours look-ahead period, satellite and NWP irradiance forecasts performed better than the persistence and climatological forecasts.

For the day-ahead period (6-48 hours), most NWP-based forecasts were more accurate than the climatological forecasts as measured by RMSE. As corroborated in previous research studies (Perez et al. 2010, Lara-Fanego et al. 2012, Mathiesen and Kleissl 2011), the error growth, i.e. RMSE vs. look-ahead time, of all NWP-based methods remained relatively modest throughout the day-ahead forecast horizon.

The hours-ahead and day-ahead composite forecasts exhibited a lower error than any individual method on average. Therefore, by combining the best available forecast information for each look-ahead time, the composite forecast provides better accuracy and foreknowledge of potential energy production than any individual method.

Although some forecasting entities do rely on more than one method to generate solar forecasts over a broad range of time scales (10 minutes to 48 hours), this project was designed to evaluate a composite forecast tool based on an optimization procedure that included several state-of-the-art forecast methods. Results of the composite forecast experiment indicate that constructing a method to capitalize on the best performing forecasting techniques at various look-ahead periods provides a tangible reduction in errors.

Lessons Learned and Next Steps

This study provided valuable lessons learned regarding the overall design of solar forecasting experiments. Major recommendations for future research efforts include:

- Universal Dataset – Target plant data was restricted by non-disclosure agreements and therefore each project team used a different dataset of solar plant data to train and validate their forecast. Differences including the total plant capacity, data availability and frequency of data output (1-second to 10-minute averages) influenced the evaluation of results and the design of the composite forecast. For example, the common data frequency interval among all team members was 10 minutes, thereby defining the minimum forecast interval used in the composite. Future research efforts should focus on obtaining a universal power production and irradiance dataset from the target plant that can be used for all aspects of the project. This would mitigate data limitations that significantly affected the minimum composite forecast interval, composite forecast design, results, and validation of individual methods.

- Real-time Forecasting – Data limitations forced the performance evaluation of the composite forecast system to be based on previous data for the calendar year 2012. Delivering results for the full year proved challenging, as the forecasting techniques are in different stages of development. The potential for real-time deployment of a composite forecast system utilizing all four methods would require a continuous and reliable flow of data from each forecast method, as well as a universal power production and irradiance dataset from the target plant itself. However, future experiments will provide tangible information that can be applied in a composite forecast design as individual methods are further refined.

The results of this project represent an important first step towards the development of an optimally integrated forecast system that can produce a more accurate forecast, for all look-ahead time frames of interest, than any single existing solar forecasting method. Therefore, method refinement of a composite forecast system, as well as planned improvements for each forecast method that will likely benefit a composite forecast are summarized as follows:

- Research based on sky imagers is currently on-going to improve its utility to capture short term ramp events. A significant improvement in sky imager forecast quality is expected with new technology by eliminating the shadowband (Urquhart et al. 2013) and the implementation of regular on-site maintenance routines. Additionally, future research efforts would benefit from the co-location of two imagers to increase the areal coverage of the images, as well as employing machine learning tools to significantly improve forecast performance and its data availability for real time forecasting. Further research is warranted to determine this method's performance on very short-term time scales.

- Satellite methods are also undergoing advancements to improve their skill under challenging conditions. Recent enhancements to satellite-based methods include the use

of multi-channel satellite data to provide pre-dawn satellite-based irradiance forecasts. Additionally, these data can be used to improve forecasts over challenging bright desert and snow covered regions. Future satellite-based forecast advancements will also include dynamic cloud models and probabilistic methodologies. Future research efforts can incorporate power production forecasts and these method enhancements.

- Much of the NWP-related effort in this project was directed toward reducing the magnitude of the systematic errors in the NWP forecasts via post processing of the NWP model outputs. There are several other major drivers of forecast performance that were not examined as part of this project, and are opportunities for future research.

  (1) Archived data available did not include global horizontal irradiance for some of the NWP system. This is a key deficiency in the application of statistical procedures to correct systematic errors in the NWP forecasts since NWP-based irradiance is the most important predictor. Future efforts can utilize modeled irradiance output from NWP forecasts in real-time or maintain an archive from the real-time data.

  (2) It is recommended that future projects make use the highest possible update frequency available from a rapid update NWP system. One near-term opportunity for this will be the use of hourly updates from the High-Resolution Rapid Refresh model, which was operated by NOAA in an experimental model during the project period but is expected to become operational in the summer of 2014.

  (3) The use of forecasts from a single NWP system results in exposure to deficiencies in a specific model's representation of future conditions. Using an ensemble of different regional models and data assimilation systems reduces the exposure to these deficiencies. Other options are available to obtain forecasts from regional NWP models, such as Weather Research and Forecast Model, Advanced Regional Prediction System, and Mesoscale Atmospheric Simulation System. AWS Truepower has configured these models to operate in rapid update mode with the assimilation of cloud information derived from satellite and Doppler radar images. The assimilation of this data has the potential to incorporate some of the short-term predictive skill found in satellite methods into the NWP methods.

  (4) Lastly, recent advances in artificial intelligence and meteorology have yielded statistical techniques such as artificial neural networks, support vector machines, analogue ensemble as well as decision-tree regression methods such as random forests and boosted regressions which have a greater ability to differentiate NWP model error regimes and apply adjustments. Only one of these approaches was experimentally applied in this project. Further exploration of the application of these methods represents additional opportunity to more precisely apply

corrections for systematic errors in NWP model forecasts and thereby improve the performance of these forecasts.

The design and implementation of the composite forecast system has provided several lessons related to the successful blending of different forecast methods. A wide variety of advanced techniques are available to combine individual forecasts into a composite forecast. A key attribute to be explored by more advanced techniques is the performance differences among the methods that are linked to differences in weather regimes. An effective exploitation of these differences would enable the weighting of the methods to vary according to their expected performance in the weather regime predicted in the forecast period under consideration. There is also an opportunity to obtain additional value from an ensemble of forecasts by constructing a probabilistic forecast. Initial steps towards the construction of a composite probabilistic forecast were advanced in this project through the use of advanced mathematical techniques. However, this capability should be further developed in future forecasting experiments.

Another potential opportunity to increase the value of a composite forecast is to formulate metrics that more precisely measure the sensitivity of the grid operator decision-making process to forecast error, and design the composite forecast method to yield the best performance for those metrics. In this project, the focus was on obtaining the lowest RMSE, but this is a very general metric that does not assess the value of the forecast for grid management decision-making very precisely. For example, the grid operator is typically much more sensitive to solar forecast error during critical periods when the solar power production is changing very rapidly (i.e. at times of large upward or downward ramps). .

Although this project focused on utility-scale solar, most of the techniques are also applicable to distributed or non-utility scale solar. However, it should be noted that the results from this project are only for one site in Nevada. The forecast performance characteristics could be substantially different for locations with different cloud and weather regime climatologies. It will be important to test these approaches for other sites with a range of cloud and weather regime climatologies.

Project Benefits

Through advancements in forecasting techniques, the cost and uncertainty associated with the integration of utility-scale solar power can be reduced. The cost-effective management of solar power production variability requires high forecast accuracy on time scales ranging from minutes to several days ahead. As solar energy production continues to increase at a rapid pace in California, the development of tools that can predict the solar energy production at multiple look-ahead time scales becomes critical to situational awareness and operational planning.

One of the significant applications of solar energy forecasts in California is to support the grid management activities of CAISO. CAISO uses solar power forecasts from individual solar generating facilities and aggregates of smaller generation resources in three specific time periods: (1) day-ahead (18-42 hours before the operating hour), (2) real-time (105 minutes before the operating hour) and (3) intra-hour (every 15 minutes for the next 2 hours). These are the primary look-ahead time periods of interest for CAISO, although 5-minute forecasting is

anticipated in the future. The forecast look-ahead horizon for this project was every 10 minutes to 48 hours, which included the intra-hour, hour-ahead and day-ahead periods relevant to CAISO operations.

The knowledge gained in this project will be directly used to improve the performance of the operational hour-ahead and day-ahead solar power generation forecasts that AWS Truepower provides to CAISO. Many of the component methods used in this project are already being employed in the ensemble forecast system used for the CAISO application. The results obtained from this project will provide the basis for improvements in two areas: (1) refinements to the component methods already being employed in the CAISO forecast production with a focus on an improved procedure to correct system errors in the NWP forecasts; and (2) a more optimal way to create the composite forecast from the individual ensemble members. It also provides a foundation for further internal research to extract additional forecast value from both of these approaches.

The complex economic and operational structure of electric systems and markets makes quantifying the benefits of forecasting difficult. However, grid integration studies have been conducted to assess the impact of integrating variable energy sources such as wind and solar energy into specific grid systems. A recent study by the National Renewable Energy Laboratory indicated that a 10% reduction in the mean absolute error of wind power production forecasts would result in an annual reduction in grid integration costs of $28 Million at a penetration level of 14 percent and an annual reduction of $52 Million at a penetration level of 24 percent (Lew et al. 2011).

While this calculation was based on wind, it is reasonable to expect that similar reductions in the error of solar forecasts would yield a similar magnitude of integration cost reductions. Furthermore, it is likely that studies have underestimated the reduction in costs because they have not considered all ways in which improved forecast information can be used to reduce costs. The focus of these studies has mostly been on the impact of more efficient day-ahead commitment of generating units. The anticipated reduction in grid operation costs with the levels of solar penetration expected in California in the near future will be realized through a more cost-effective decision-making process on multiple time scales ranging from minutes to hours ahead. The process will include decisions for regulation and load following purposes, in addition to day-ahead commitment of generation resources.

Therefore, advancements in solar forecast performance will result in interrelated economic and social benefits including (1) a reduction in the cost of integrating solar power into grid planning and operations, and (2) facilitation of higher solar penetration levels at a specified cost and reliability. Improvements in forecast performance can result in lower cost electricity and the associated societal benefits from a reduction in greenhouse gas emissions.

CHAPTER 1:
Introduction

1.1 Background and Overview

Solar power production can fluctuate rapidly between days, hours and even minutes as a result of atmospheric conditions, including clouds and aerosols, significantly affecting grid operations and management. The goal of this project was to facilitate grid integration of utility-scale solar energy through advancements in forecasting techniques of variable power output.

Forecasting methods targeted for specific look-ahead periods, such as hour-ahead or day-ahead, have been developed and are in operational use by grid operators. Previous research and operational experience has indicated that there is no single forecast method that is optimal for all relevant look-ahead periods. Four approaches that have demonstrated considerable skill for this range of look-ahead times include: (1) sky imager - cloud motion vector, (2) satellite - cloud motion vector, (3) rapid update NWP, and (4) standard cycle regional and global NWP. However, an approach that provides the highest possible accuracy over a broad range of time scales by utilizing different forecast methods has not been implemented or evaluated. This project configured, demonstrated and validated a composite forecast tool that used the best of each method for the targeted look-ahead period to produce a solar forecast. The forecast look-ahead window was defined as 0 to 48 hours, ranging from forecasts at very short time scales of minutes (intra-hour) to hours-ahead, and day-ahead.

1.1.1 Challenges of Integrating Solar-Generated Electricity

Passing clouds can significantly influence the power production at any given solar plant location in real-time. Grid operations that have traditionally relied on dispatchable, steady power production are challenged by utility-scale solar energy plants that can experience steep power ramps (large excursions of power production over short time scales) caused by the shadows of relatively fast moving, boundary layer (low-level) cumulus, stratocumulus and altocumulus clouds (see Figure 1). Forecast information about future potential solar energy production allows grid operators to capture and manage the variability of solar power generation (see Figure 2), so it can be considered in system operations and planning, such as load following and unit commitment.

Figure 1: Schematic representation of common types of clouds

Source: http://eo.ucar.edu/webweather/cloud3.html
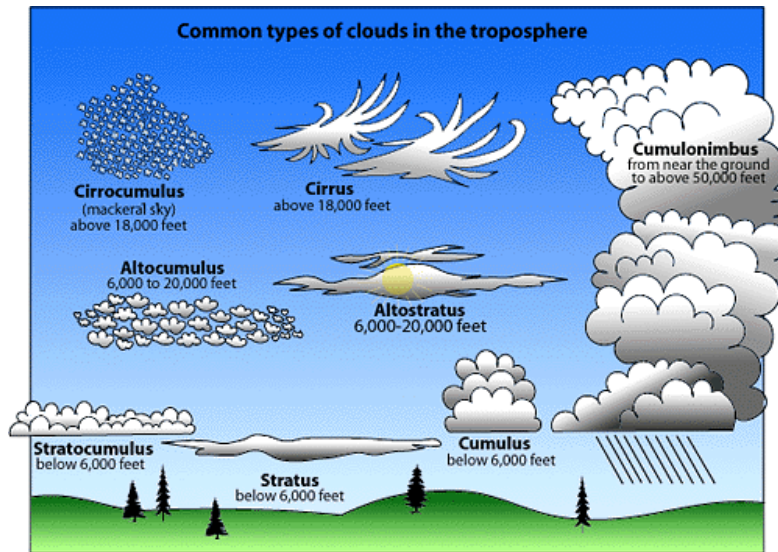
Figure 2: A typical daily pattern of direct and diffuse horizontal irradiance (blue and red lines, respectively) measured by a rotating shadowband radiometer.

Source: http://eo.ucar.edu/webweather/cloud3.html
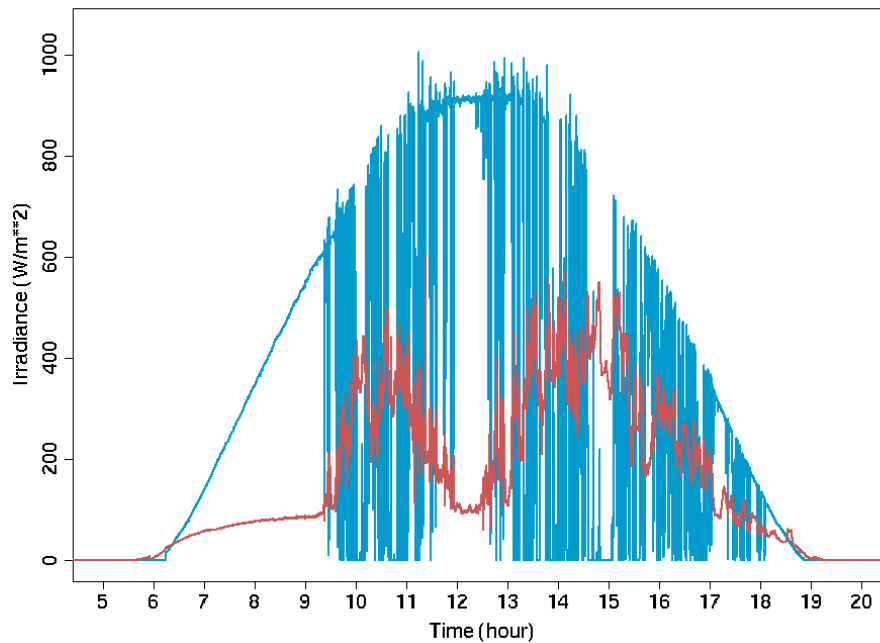
Since PV systems respond almost instantly to changes in irradiance, similar patterns are seen in PV output.

1.1.2   Advanced Short-term Solar Resource Forecasting

Advanced meteorological modeling applications incorporate various technologies and data sources including ground-based sun-tracking cameras, whole sky imagers, visible and infrared satellite imagery and NWP models. Clouds can be detected and characterized to predict solar irradiance using ground-based observations or satellite imagery. Forecasts can be made from these observations by knowing the "current" position of the clouds and inferring on their velocity based on recent history of the cloud's positions. Another common approach to generating solar irradiance forecasts is through the use of computer models such as NWP models. These numerical models solve the equations of fluid dynamics together with a full suite of physics packages. These NWP models therefore simulate the evolution of atmospheric conditions such as temperature, precipitation, winds, etc. and more importantly, cloud cover fraction and solar irradiance.

Solar resource and power production forecasting methodologies perform fairly well at time scales of less than one hour up to a few days in advance. As described in Table 1, sky imagers, satellite-cloud motion vector and NWP approaches each have advantages and disadvantages when applied as stand-alone forecasting applications. Satellite-cloud vector forecasting performs best at time scales of 0.5-5 hours, whereas NWP models perform best at time scales greater than 5 hours ahead (Perez et al. 2010). Neither approach is currently sufficient to capture events that occur on small time scales and affect power production and forecast accuracy.

Table 1: Forecasting Methods

Method	Advantages	Disadvantages
Sky imagery: intra-hour forecast approach (0 - 30 minutes)	Very detailed information about the extent, structure, and motion of existing clouds in the vicinity of the solar generation facility at the time the forecast is made.	Does not account for cloud development and dissipation or significant changes in cloud geometry; extrapolation of cloud patterns is limited by the field of view of the sky imager.
Satellite-cloud vector approach (0.5 - 5 hours)	Much larger spatial scale of cloud patterns can be detected; motion of the cloud field can be projected forward over longer time periods.	Spatial resolution of these images is much less than sky imagers; small clouds and small structures in larger scale cloud features not detected.
Numerical Weather Prediction (NWP) Forecasts (5 – 48+ hours)	Ability to simulate cloud development, dissipation, and structural changes as well as cloud translation in the flow field.	Typical size of model grid cells (~ 5 to 10 km+) prohibits the representation of small-scale features such as cumulus clouds; no accurate representation of initial cloud patterns.

The best forecast performance for 0-1 hour look-ahead period (0-1 hours ahead) is typically achieved by methods that can diagnose small-scale details of current cloud patterns, their recent evolution and project the trend forward to obtain a short-term forecast.

Ground-based, sun-tracking cameras and sky imagers define the local cloud structure and motion. Methods to exploit the predictive value of the information obtained from sky imagers have been developed and evaluated (e.g. Chow et al. 2011 and Yang et al. 2014) by project team member, University of California - San Diego. This approach is limited by two fundamental factors: (1) the field of view of the sky imager and (2) the rate of development and dissipation of the type of clouds being observed by the sky imager since these effects are not included in the current formulations of the cloud vector methods. For a single sky imager, the first factor is probably the most significant and typically limits the prediction horizon to 20 minutes or less for the typical range of cloud speeds. The use of multiple sky imagers can overcome this problem to some extent and in that case, the most significant limitation generally shifts to the second factor. The useful look-ahead time from this approach is generally 30 minutes or less with the greatest value being obtained in the 0 to 20 intra-hour, minute-ahead time frame.

The optimal method for the 0-5 hours-ahead time frame is typically a satellite-based cloud vector approach (Perez et al. 2010). The approach has been extensively used and refined by project team member, Clean Power Research. This approach is conceptually similar to the cloud vector approach used with the sky imager data. However, the images provided by the satellite-based sensors have a much larger geographic field of view, but their spatial and temporal resolution is somewhat lower than the sky imager data. The typical spatial resolution for visible satellite image data is about 1 kilometer. Therefore, the satellite-based cloud vector approach is not limited by field of view considerations, but is limited by its resolution (inability to see small cloud features) and the rates of development and dissipation of cloud features. Features that are significantly larger than 1 km, relatively stable in shape and size, and move at a constant velocity can frequently be predicted for many hours. However, it is usually difficult to predict rapidly evolving multi-layer cloud features for more than an hour with this approach.

Another method that can be useful for the 1-5 hour, look-ahead period is a rapid update NWP model. Project team member, MESO, Inc. has considerable experience with the development and application of rapid update NWP models. The NWP approach is based on a form of computational fluid dynamics in which Navier-Stokes equations are used to calculate the future state of an atmospheric volume given a specification of the current (initial) state. The NWP approach has been used with considerable success for day-ahead and longer weather prediction intervals.

The rapid update version of NWP employs a more detailed, higher resolution NWP model using an update cycle of every 1 or 2 hours. The rapid update NWP model predicts cloud development and dissipation as well as simple and complex cloud movement patterns. The NWP method is typically limited by the sparseness of data available to adequately specify the cloud structures (e.g. cloud top, cloud base, liquid and frozen water distribution etc.) and other basic atmospheric variables (e.g. temperature, relative humidity, wind speed and direction etc.) for the initialization of the forecast. The demanding data requirements usually limit the

practical value of this approach for short-range cloud forecasting. A previous study by Perez et al. (2010) showed that the satellite-based cloud vector approach outperforms the NWP method for look-ahead times under 5-6 hours. However, the increased update frequency for rapid update NWP systems suggests that the value of this tool may improve model-based, very short-range forecasts.

Solar forecasts produced many hours to days in advance (6 to 48 hours in this project) typically employ the standard cycle 6-hour or 12-hour update NWP forecasts. These forecasts are produced by the operational government forecast centers such as the National Weather Service's (NWS) National Centers for Environmental Prediction (NCEP). Typically, an ensemble of NWP forecasts is used for the day-ahead and longer forecasts. However, forecasts of most variables produced by NWP models typically have significant systematic errors. Therefore, the forecasts usually are adjusted with some form of statistical model using data from the forecast site. This is a process known as model output statistics (MOS).

1.1.3 Composite Forecast Experiment

The composite forecast experiment integrates four solar forecasting approaches (sky imager-based cloud motion vector, satellite-based cloud motion vector, rapid refresh and standard cycle NWP) to configure, demonstrate and validate an advanced multiple-time scale, state-of-the-art solar forecasting tool. Solar forecast performance was evaluated at multiple time scales relevant to grid operations and reserve planning using three different look-ahead periods:

- intra-hour (10 minutes to 30 minutes)

- hours-ahead (0.5-5 hours)

- day-ahead (6-48 hours) forecast horizon

The individual performance of each forecast method during the look-ahead period of interest was evaluated relative to each other and standard reference forecasts (e.g. persistence or climatology). The best performing methods were then integrated to produce a first composite forecast tool.

1.2   Project Goals

The goal of this project was to reduce the cost of integrating utility-scale solar plants into the electric system by improving forecasting techniques used to predict solar energy production on short look-ahead time scales. Four different forecasting methods with strengths and weaknesses that varied by look-ahead interval were evaluated and integrated into a seamless composite forecast system to provide timely and accurate foreknowledge of intra-hour through day-ahead solar-based generation for use in grid system operations. Thus, the composite forecast system was able to take advantage of the strengths of each method in constructing an optimal integrated forecast for each look-ahead period. Increasing penetrations of variable energy sources will drive grid operations to consider these sources in load following, unit commitment and the system operation decision-making processes. Improving the accuracy of solar forecasting tools and methods will increase the utility of this information to grid operators.

CAISO was a key project partner that provided guidance on the desired specifications of solar forecasts (update frequency, time resolution, look-ahead period etc.). The CAISO guidance allowed the project team to maximize the contribution of composite forecast experiment to the cost-effective management of solar power production variability on the electric systems operated by CAISO and other balancing authorities.

1.3   Report Structure

Section 2 of this report outlines the project approach, including task organization and a description of each major effort as presented in Sections 3, 4, 5 and 6. Section 7 provides a summary of conclusions, outlines technical recommendations and findings for future research opportunities, and includes a discussion of the public benefits provided by this research effort. Section 8 outlines the knowledge transfer activities undertaken as part of this project.

CHAPTER 2:
Project Approach

The objective of this project was to configure, demonstrate and validate a solar forecasting tool that optimally integrates several forecast methods to achieve the best forecast performance within the targeted forecast look-ahead periods. Four approaches that have demonstrated considerable skills were considered in this research project include: (1) total sky imagers, (2) satellite - cloud motion vector, (3) rapid update NWP, and (4) standard cycle regional and global NWP.

The project was divided into four technical tasks to develop and test a composite forecast system. Copper Mountain, a utility-scale photovoltaic power plant provided the basis for data collection and validation of the individual methods and the composite forecast experiment.

2.1 Atmospheric and Plant Data Acquisition

The goal of this task was to gather, validate and quality-control meteorological, forecast and target plant data necessary to produce each type of forecast and evaluate its performance. A variety of data sources are required as input for each forecast method. Activites included:

- Collection and archiving of on-site data from sky imagers,

- Quality control of sky imager data, and determination of sample days for training and testing

- Collection and archiving of satellite imagery, and specific GHI data from Copper Mountain weather stations

- Collection and archiving measurement data from a diverse mix of atmospheric sensors deployed by public and private sector entities.

- Collection and archiving of operational large-scale NWP model output and forecast datasets including GFS, NAM, and RAP models.

Many sources of input data used by the forecast methods are publically available, such as GOES satellite data or the NCEP model output. However, each forecast method also requires observed on-site irradiance (GHI) and operational plant production data (MW) for model training and validation purposes. This data is considered proprietary and attempts to gain access to a universal data set for use amongst the team members were not successful. Although each dataset was obtained through various means and corresponds to different averaging periods, the analysis indicated that the data sets likely originate from the same underlying source. However, it required modifications to the composite forecast methodology and evaluations of individual forecast results, which are summarized in the relevant sections of this report.

2.2   Solar Forecast System Setup

Each forecast method employed in the composite forecast system generated an individual solar irradiance forecast and its performance was evaluated for its respective look-ahead time frame. Activities included:

- Documented individual forecast methodologies.

- Generated individual forecasts from each method.

- Evaluated individual forecast performance using standard forecast metrics.

- Compared against performance of individual methods persistence and climatological forecasts.

During this process, it was determined that a 10-minute average was necessary to serve as the baseline for evaluation due to data access limitations. All team members had access or could generate a 10-minute average for comparative purposes.

2.3   Design and Evaluate Composite Forecast System

Forecasts for a 1-year period extending from January 1, 2012 to December 31, 2012 were generated by each method for use in the composite forecast system. The composite system was designed and then evaluated. Activities included:

- Developed and documented creation of a "short-term" ensemble designed for the 0 to 6 hour-ahead forecasts with a 10-minute time resolution and an update frequency of 10 minutes.

- Developed and documented creation of a day-ahead ensemble designed for the 0 to 48 hour-ahead forecasts with a 1-hour time resolution with an update frequency of 1-hour.

- Determine the proper weighting of each individual forecast method for the relevant look-ahead period.

- Process and archive composite forecast results for 2012.

- Evaluated composite performance using standard forecast metrics.

- Compared performance of the composite forecast against persistence and climatological forecasts.

2.4   Knowledge Transfer Plan

The purpose of this task is to disseminate the knowledge gained, experimental results and lessons learned to the intended users of project results including industry stakeholders, utility operators, the California Independent System Operator (CAISO), as well as the domestic and international public, private and academic research and forecasting community at large. Activities were documents in Monthly Progress report. Activities included:

- Prepared a Technology Transfer Plan.

- Results of this project were orally presented at two internationally renowned conferences:

    o 3rd International Workshop on Integration of Solar Power into Power Systems on October, 2013 in London

    o The American Meteorological Society's (AMS) Annual Meeting February 2-6, 2014, in Atlanta, Georgia

- Provided regular briefings to CAISO regarding the design of the composite forecast system and information about its performance characteristics. The team collectively explored how to achieve maximum value for the grid integration of solar-based generation from these forecasts.

The plan anticipates the participation of the project team in a Lessons Learned Information Workshop to be hosted by the Commission in the future.

CHAPTER 3:
Atmospheric and Plant Data Acquisition

3.1   Introduction

Copper Mountain Solar 1, a 58 megawatt (MW) utility-scale photovoltaic power plant in Nevada, provided the basis for data collection and validation of the individual methods and the composite forecast experiment. Copper Mountain was chosen as the forecast target and evaluation site because the power generated at this plant is fed into the CAISO grid.

A variety of data sources are required as input for each forecast method utilized in this project, including the collection of on-site data, such as sky imagery, solar irradiance and power production. Many sources of data required for this project were publically available. However, the use of proprietary of plant data was restricted by non-disclosure agreements, prohibiting data sharing or use amongst the project partners. Each participant was able to gain access through its existing agreements to plant power production data. The use and validation procedures employed as a result of restrictions on data for each individual method is described in subsequent sections of this report.

This section provides an overview of the data required for each methodology, its collection, quality control efforts and challenges for the individual forecast methods.

3.2   Operational Plant Data

The Copper Mountain Solar 1 power plant, located in Henderson, Nevada is owned and operated by Sempra US Gas and Power (Figure 3). The solar plant, which totals 58 MW, actually comprises two separate projects, a 48 MW project and a 10 MW project. In this study, only data from the 48 MW project was used. The Copper Mountain Solar 1 plant uses cadmium telluride thin film PV panels for 96 - 500 kilowatt inverters. The panels cover approximately 1.3 km^2.

Fifteen National Renewable Energy Laboratory calibrated panel modules provide plane-of-array irradiance and five weather stations provide standard meteorological measurements (i.e. wind speed, temperature, relative humidity) including plane-of-array irradiance and GHI at the target site. These measurements served as reference stations for validation and testing.

Occasional on-site maintenance and daily data screening from the TSI units is critical to data assimilation for the modeling efforts. Power plant personnel were trained to provide periodic maintenance by cleaning and polishing the mirror of the sky imager. Maintenance and cleaning of the mirror was recommended for every two weeks.

Figure 3: 48 MW Copper Mountain Solar 1 power plant.
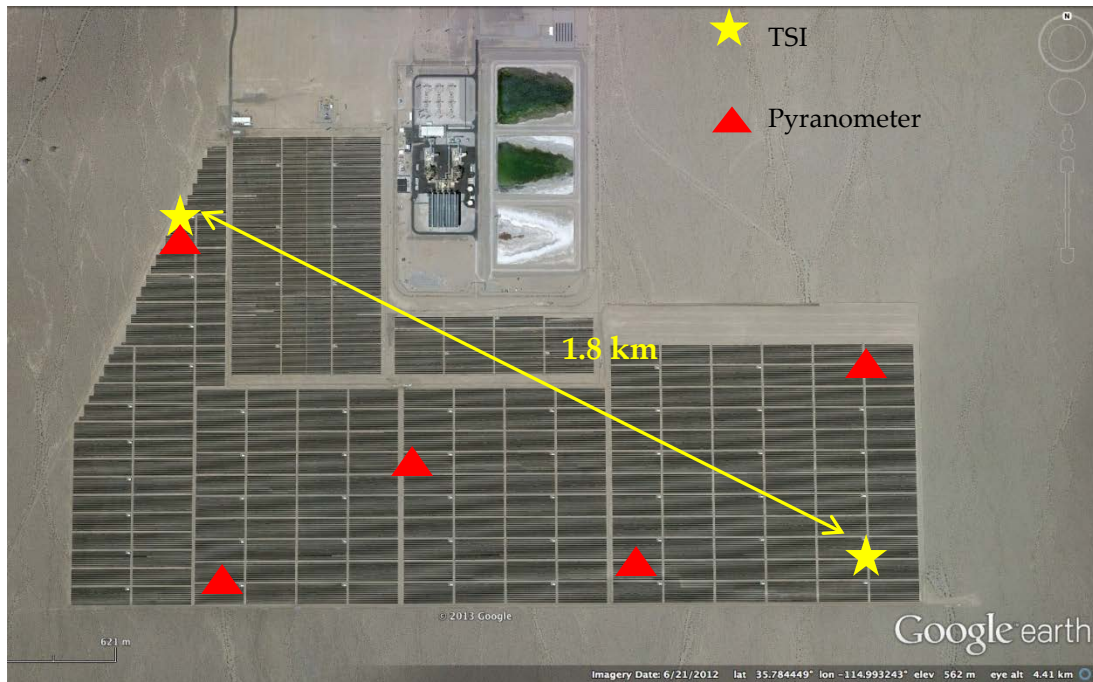
Photo Credit: Google Earth. Data Source: University of California, San Diego

3.3 Input Data by Forecast Method

3.3.1 Total Sky Imager Forecast: Target Site Data Collection

The sky imager method was identified as a forecasting method that may produce viable results in the intra-hour look-ahead time frame with generally the greatest value being obtained in the 0 to 20 minutes-ahead time frame. Sky imagers are used to map the cloud field and obtain a projection of current and future cloud shadows crossing a utility-scale solar system array. The Copper Mountain Solar 1 PV plant in Henderson, Nevada, was instrumented to capture the fine scale temporal and spatial variability of the cloud-light environment.

Two TSI440s were installed at the target site (Figure 4) in the summer of 2011 and were leveraged for use in this project to validate the sky imager forecast methodology in a utility-scale environment. The initial installation and maintenance of the TSI units was supported by funding from the National Renewable Energy Laboratory.[1] Additionally, the initial setup of the forecast algorithms and data analysis was supported by funding from a separate Energy Commission contract.[2] These two awards provided the foundational instrumentation to support this project effort. The current contract provided for continued operation and advancements in the processing of additional sky imager data. The TSIs are spaced approximately 1.8 km apart. TSI images are acquired and archived every thirty seconds from the two units installed at the target plant (TSI-1 and TSI-2).

---

[1] Subcontract AXL-1-40357-01.

[2] CEC Contract # 500-10-043.

Figure 4: Total Sky Imager at Copper Mountain

Photo Credit: University of California, San Diego

Quality control was performed on the acquired TSI image data. This process consists of visually inspecting selected images for each day data is available. Data is excluded from consideration during the quality control process for a variety of reasons. For example, when the shadowband (which is designed to shield the camera optics from the sun) becomes misaligned the data may not be useable and it is excluded. Poor image quality caused by excessive soiling of the TSI's mirror can also result in data exclusion. Clear days were also identified and excluded from consideration as TSI imagery on clear days does not provide useful data for model forecasts. Significantly overcast or rainy days are similarly excluded from further analysis.

Table 2 provides the percentages of the excluded data from 2012 and the cause for exclusion. The data availability for each day in the period of record (1/1/2012 – 12/31/2012) is classified as "good", "not good" or "no data'. Days classified as "not good" were primarily due to soiling of a TSI (e.g. from dust, precipitation) and occasional misalignment issues with the shadowband on a TSI. In 2012, daily data for the TSI was available for 299 days out of the year (i.e. 18% missing data) due to temporary errors in irradiance or power measurements, usually due to database or sensor outages. Of the 299 days, 62 were considered viable candidates for model testing. A large percentage of days were excluded due to clear sky weather conditions common to the desert environment (38%); most non-weather exclusions were attributed to dirty images (22%).

Table 2: TSI Data Availability for Model Assimilation in 2012

Data Availability for Project Analysis 1/1/2012 – 12/31/2012	Number of Sample Days	Percentage of All Days
Total Number of Days Data was Available	299	82%
Total Days Excluded from Project Analysis		
"Not Good" – Dirty Images	80	22%
"Not Good" – Shadowband Misalignment	12	3%
"Good" – Clear Day	139	38%
"Good" – Rainy or Overcast	6	2%
Total TSI Days Available for Analysis	62	17%

For basic validation of cloud detection from sky imagery, the TSI shadow fraction over the plant was compared to on-site irradiance measurements for a sequence of days with adequate data recovery and the preferred atmospheric conditions (variable cloud cover). The irradiance shadow fraction is defined as the fraction of time when the clear sky index of the ground measurements was less than 0.7. Table 3 shows that TSI shadow fraction (i.e. fraction of the plant covered by a cloud shadow) has some consistency with on-site irradiance measurements. Note that TSI shadow fraction represents an instantaneous, whole sky depiction of the cloud state; whereas the irradiance shadow fraction uses point measurements to estimate the light-sky conditions. Exact agreement is not expected as the sky imager and the ground stations sample different areas of the sky. Larger disagreements on July 13 and 16 were related to erroneous cloud detection in the circumsolar region of the TSI images, where strong forward-scattering by aerosols causes a spectral light signature that emulates that of clouds.

Table 3: Comparison of Irradiance and Sky Imager Shadow Fraction.

Day	TSI Data Availability [%]	TSI Shadow Fraction [%]	Irradiance Shadow Fraction [%]
July 12 2012	99.8	80.5	79.4
July 13 2012	99.9	30.8	73.2
July 14 2012	99.6	52.7	36.3
July 15 2012	99.7	67.6	24.7
July 16 2012	99.9	41.8	9.9
July 17 2012	99.9	17.3	1.0

3.3.2 Satellite-Based Forecast: Geostationary Operational Environmental Satellite (GOES)

In support of the satellite-based forecast method, GOES imagery are acquired via FTP from the National Oceanographic and Atmospheric Administration (NOAA) data servers and archived. The GOES visible satellite images record visible light (from the sun) that is reflected by clouds, land, oceans, snow, and ice (Figure 5). GOES data are sampled half hourly at one kilometer resolution over the project area. Raw GOES satellite imagery has been archived every 30 minutes at a one kilometer resolution for this project since September 2011.

Figure 5: GOES visible satellite image for January 12th, 2014 at 20 UTC.

Source: http://www.goes.noaa.gov
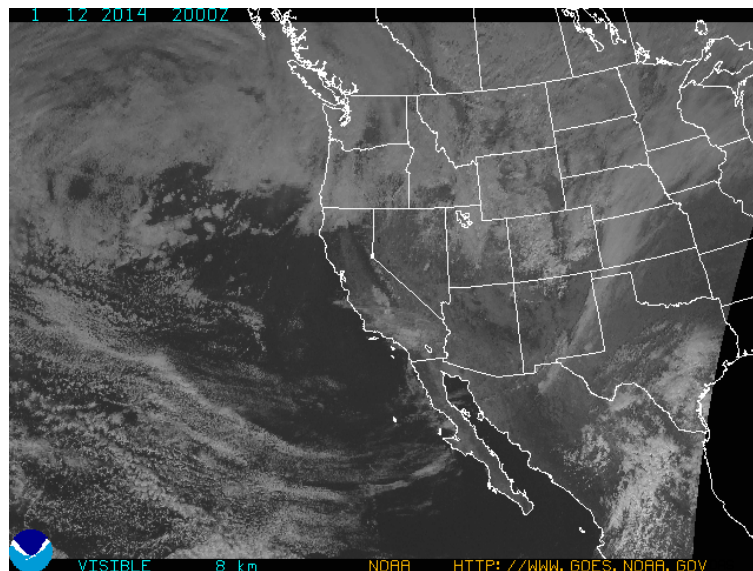
The GOES 1 km visible satellite images were used in the SolarAnywhere model to generate high resolution forecasts. A 1 km wide pixel centered on the Copper Mountain Solar 1 site is selected to represent the downwelling surface irradiance conditions at the solar plant (Figure 6). The SolarAnywhere model used to derive surface irradiance is "self-calibrating" and uses the recent dynamic range of the GOES satellite data for model calibration needs (Perez et al. 2002). Raw GOES satellite data is re-calibrated each year to account for known satellite data bias/drift issues as reported by NOAA. This helps insure year-to-year continuity in the visible satellite data record. Monthly satellite data recovery is shown in Table 4.

Site-specific GHI data from the five Copper Mountain weather stations were visually inspected, quality controlled and then averaged into 10-minute and half-hourly segments to quality control the satellite images and to train the forecasting system. Individual GOES satellite images were quality controlled via automated processing and checked for missing and suspect data (noise, missing scan lines, etc.) before being incorporated into the Solar Anywhere model.

Figure 6: SolarAnywhere® 1 km grid location chosen for the Copper Mountain satellite-based forecasts.
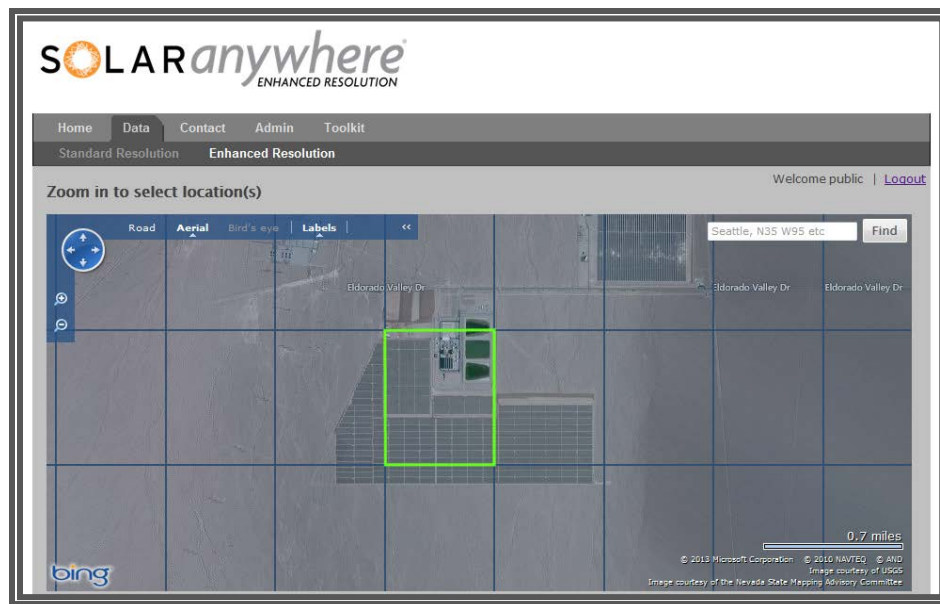
Source: Clean Power Research. Image courtesy of USGS and Nevada State Mapping

Table 4: Monthly Data recovery (%) from GOES visible satellite imagery

	Jan	Feb	March	April	June	July	Aug	Sept	Oct	Nov	Dec
2011								99.72	99.73	99.72	99.72
2012	99.73	99.71	98.39	99.72	99.73	99.72	99.73	99.73	99.73	99.72	99.73

3.3.3 Numerical Weather Prediction Model Data Archive

The NWP method consists of two components: (1) output from the standard cycle (every 6 or 12 hours) NWP models; and (2) mesoscale NWP models that operate in rapid update data assimilation cycle (every 1 or 2 hours). The first component is focused on the day-ahead and longer portions of the forecast period while the second component is designed to provide information for the intra-hour (2 to 15 hours) look-ahead period. Both components employ a MOS procedure that is designed to reduce systematic errors in the forecast derived from each NWP model.

Two types of data are used in the Numerical Weather Prediction (NWP) portion of the forecast system: (1) large-scale NWP analyses (i.e. model initialization data) and forecast datasets; and (2) measurement data from a diverse mix of atmospheric sensors deployed by public and private sector entities. Operational model output data is archived via NOAA's National Operational Model Archive and Distribution System (NOMADS). These data include the output from NCEP's GFS, NAM, Rapid Update F Cycle (RUC) and RAP models.

Data collection began in September 2011 for the GFS and NAM. The GFS and NAM model runs are executed four times per day at 0000, 0600, 1200 and 1800 UTC. The RAP became operational on May 1, 2012 and prior to May 1, 2012, the RUC data was archived. The RAP and RUC models are executed hourly, however data was archived at two hour intervals (i.e. every other cycle). Although there are differences in the physics and data assimilation algorithms included in the RUC and RAP models, they were treated as one system for the purposes of the statistical model employed in the composite system. The RUC and RAP data was used to provide input to the MOS algorithms and the forecast results have been identified for the purposes of this report, as simply the RAP model.

Infrared and visible satellite imagery from the NOAA GOES satellite was collected since September 2011. Infrared images were archived at a rate of one image per hour for 24 hours a day at 4 km grid resolution. Visible satellite imagery was been collected and archived at a rate of four images per hour (i.e. every 15 minutes) for 20 hours a day on a 2 km grid resolution (native resolution is 1 km). The data is archived at a resolution of 2 km in order to conserve storage space. The NWP resolution is coarser than 2 km and therefore the extra resolution does not provide value for the NWP application. The infrared satellite image data is used to add moisture information to the initial state of the rapid update NWP models to improve the representation of cloud location, coverage and optical depth.

Atmospheric measurement data from public and private sources (such as from weather balloons, wind profilers, surface weather stations etc.) and a review of local surface irradiance stations near the target plant was completed. Quality-controlled sets of many types of publically-available meteorological data (e.g. weather balloons, wind profilers, surface weather stations, etc.) are available in near real-time on demand through the Meteorological Assimilation Data Ingest System (MADIS).[3] However, satellite image data and NOAA radar data are not available and therefore a separate archive of this data was maintained. The

---

[3] See: http://madis.noaa.gov

publically available data on MADIS, along with the satellite image data, for the initialization of the rapid update NWP models.

Tables 5, 6, 7, and 8 provide a summary of the four types of NWP data that was archived for use in this project. The NWP data available from the government do not include forecasted GHI values from some of the NWP models. This is a key deficiency since the NWP forecast of GHI is a key predictor of the actual GHI. Actual forecasted GHI values were available in the archive of the NAM model data, but were not available for the GFS, RUC or RAP models. Therefore, an algorithm was developed to estimate the forecasted GHI for the GFS, RUC and RAP models from the variables that were available in the data archives. The algorithm used the vertical profile of relative humidity for each model which was available at 50-millibar intervals for the GFS model and 25-millibar intervals of pressure for the RUC and RAP models.

Table 5: NWS GFS Model Data Summary

Data Type (Degree Grid Point Data)	Period of Record Begins	Forecast Cycles Available	Data Recovery (%)
0.5	Sep 2011	120/120	100%
0.5	Oct 2011	124/124	100%
0.5	Nov 2011	120/120	100%
0.5	Dec 2011	124/124	100%
0.5	Jan 2012	124/124	100%
0.5	Feb 2012	116/116	100%
0.5	Mar 2012	124/124	100%
0.5	Apr 2012	116/120	97%
0.5	May 2012	118/124	96%
0.5	Jun 2012	119/120	99%
0.5	Jul 2012	124/124	100%
0.5	Aug 2012	124/124	100%
0.5	Sep 2012	120/120	100%
0.5	Oct 2012	124/124	100%
0.5	Nov 2012	120/120	100%
0.5	Dec 2012	124/124	100%
0.5	Jan 2013	100/124	81%

Table 6: NWS NAM Model Data Summary

Grid Type	Period of Record	Forecast Cycles Available	Data Recovery (%)
12 km	Sep 2011	120/120	100%
12 km	Oct 2011	124/124	100%
12 km	Nov 2011	120/120	100%
12 km	Dec 2011	124/124	100%
12 km	Jan 2012	124/124	100%
12 km	Feb 2012	116/116	100%
12 km	Mar 2012	124/124	100%
12 km	Apr 2012	116/120	97%
12 km	May 2012	118/124	96%
12 km	Jun 2012	119/120	99%
12 km	Jul 2012	124/124	100%
12 km	Aug 2012	124/124	100%
12 km	Sep 2012	120/120	100%
12 km	Oct 2012	124/124	100%
12 km	Nov 2012	120/120	100%
12 km	Dec 2012	124/124	100%
12 km	Jan 2013	100/124	81%

Table 7: NWS Rapid Update Cycle (RUC) Model Data Summary

Grid Type	Period of Record	Forecast Cycles Available	Data Recovery (%)
13 km	Jan 2012	345/372	93%
13 km	Feb 2012	345/348	99%
13 km	Mar 2012	312/372	84%
13 km	Apr 2012	360/360	100%

Table 8: NWS Rapid Refresh (RAP) Model Data Summary

Grid Type	Period of Record	Forecast Cycles Available	Data Recovery (%)
13 km	May 2012	275/372	74%
13 km	Jun 2012	357/360	99%
13 km	Jul 2012	338/372	91%
13 km	Aug 2012	328/372	88%
13 km	Sep 2012	343/360	95%
13 km	Oct 2012	372/372	100%
13 km	Nov 2012	360/360	100%
13 km	Dec 2012	372/372	100%
13 km	Jan 2013	285/372	76%

3.4   Data Sources

Many sources of input data used by the forecast methods are publically available, such as GOES satellite data or the standard NWP model output. However, each forecast method also requires observed on-site irradiance and operational plant production data (MW) for model training and validation purposes. This data is considered proprietary and attempts to gain access to a universal data set for this project were not successful. Table 9 shows the difference in data characteristics for each method in this study.

In order to confirm a common baseline, datasets used for model training and validation purposes were evaluated for deviations and anomalies that may indicate major differences amongst the datasets available to each team member. It is critical that each method utilizes the same underlying source, such as the observed irradiance and power data, used in the training and validation of the individual and composite forecasts. Histograms of normalized power generation and irradiance data utilized for each forecast method were evaluated. Although each dataset was obtained through various means and corresponds to different averaging periods, the analysis indicated that the data sets likely originate from the same underlying source. Therefore, the project team had confidence that each method could generate forecasts using their unique observed irradiance and power datasets and that those forecasts would remain consistent for each method.

Table 9: Characteristics of observed irradiance and actual power generation data from the Copper Mountain Solar 1 site for each method in the study.

	NWP	**Sky Imager**	**Satellite**
Parameters	GHI, plane-of-array irradiance and power generation	GHI, plane-of-array irradiance and power generation	GHI and power generation
Averaging interval of raw data	10 minutes	1 second	1 minute

Chapter 4:
Individual Forecast Method Description

The objective of this project was to assemble and evaluate a solar irradiance and power generation forecast system that optimally integrates several forecast methods to achieve the best forecast performance within the targeted forecast look-ahead periods. Previous research and operational experience has indicated that there is no single forecast method that is optimal for all relevant look-ahead periods. Four approaches that have demonstrated considerable skill for this range of look-ahead times include: (1) total sky imager - cloud motion vector, (2) satellite - cloud motion vector, (3) rapid update NWP, and (4) standard cycle regional and global NWP. The NWP subsystem is composed of elements (3) and (4) and is denoted by the gray shading in Figure 7.

Each of these methods produces a forecast of solar irradiance and power generation. The NWP methods also produce a forecast of other relevant meteorological variables such as temperature. Each method produces forecasts for the look-ahead period for which it is considered to be relevant. Thus the sky imager – cloud vector approach produces forecasts for less than a 1-hour look-ahead period while the standard cycle NWP will produce forecasts for the full 0 to 48-hour forecast period that is considered in this project. The following sections summarize the technical characteristics of each approach, including their strengths and weaknesses.

Figure 7: Composite Forecast Method
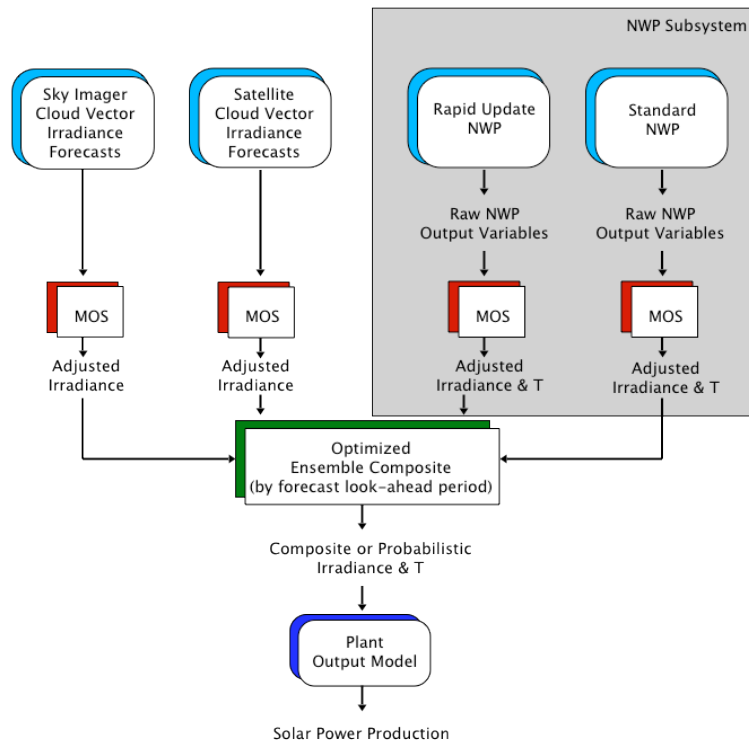
Source: AWS Truepower

4.1 Total Sky Imager Method - Cloud Motion Vector

4.1.1 Intra-Hour Experimental Setup

The method used to generate total sky imager forecasts for the Copper Mountain case study follows that of Chow et al. (2011). The forecast procedure is outlined in the flow chart in Figure 8. The procedure is broken up into two steps - one that relies on sky imager data and one that is designed for the power plant being studied. After a new image is collected, image specific masks (for the sun, shadowband, etc.) are applied and the image pixels are georeferenced to zenith and azimuth angle pairs. Using cloud altitude computed from nearby METAR stations at KLAS, KHND, KVGT, and KLSV, image pixels are projected to a planar latitude-longitude grid at a specific altitude which is termed the 'cloudmap'. The cloud height at the solar plant is chosen based on a weighting scheme that takes into account distance the solar plant and the METAR station (the closer, the larger the weight) and current cloud observations (the more cloudy, the larger the weight). From that station the lowest layer of scattered or denser clouds is chosen as the cloud altitude. Of the two TSIs installed at the plant, only the northwestern unit (TSI1) was used in this case study to generate cloudmaps for forecasting, because it was typically situated in the upwind direction. Cloud detection, cloud motion, and forecast generation are described in the following sections.

Figure 8: Flow chart showing the basic operations for constructing the power forecast using total sky imagery.
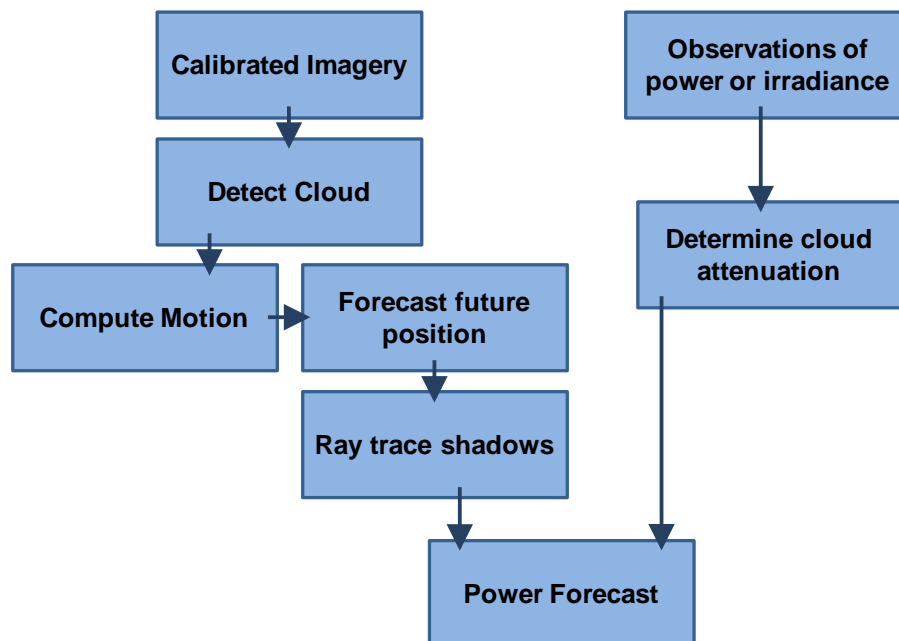
Source: AWS Truepower

4.1.1.2 Cloud Detection

The ratio of the red channel to the blue channel from the raw imagery, better known as the red-blue ratio (RBR) is the primary metric for cloud detection. The RBR of a cloud in a digital image

is close to one whereas for clear areas the ratio is much less than one (Figure 9). After characterizing the typical RBR of clear pixels through a clear sky library (CSL), a threshold for RBR excess over the clear sky expected value is set to distinguish between cloudy and clear states. The CSL tabulates clear sky RBR as a function of pixel zenith angle (the angle a pixel makes with the zenith) and sun-pixel angle (angle a pixel makes with the sun). Cloud detection yields a binary mapping of sky conditions.

Figure 9: Cloud detection process through conversion of the raw image into a red-blue ratio map adjusted for the time-of-day through a clear sky library and finally converted to a binary image (cloud, no cloud).
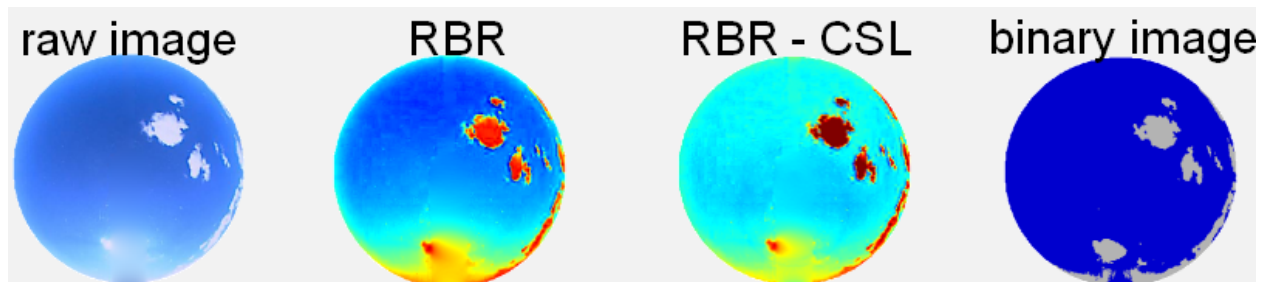
Source: University of California-San Diego

4.1.1.3 Cloud Velocity

The speed and direction of clouds is determined by analyzing the change in cloud position between consecutive images. A change in cloud position is detected using a normalized cross correlation procedure. The process begins by first segmenting one image into small tiles, and then each of these tiles is cross correlated with the other image. The tiling procedure and the corresponding search area are illustrated in Figure 10.

Figure 10: Illustration of the normalized cross correlation method used to compute inter-image cloud motions.

a)

b)
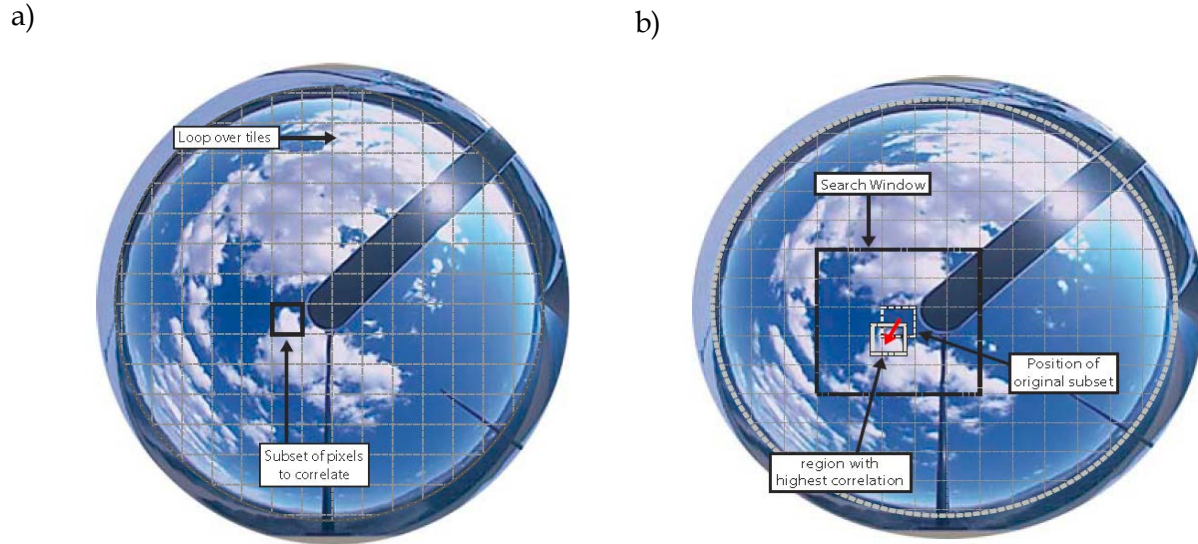
The image (a) at t_0-30 sec is broken into small tiles and each of these tiles cross-correlated with the corresponding search window in (b) the image at t_0. Note that while the process is illustrated using raw images in the coordinates of the imaging system, for the implementation images transformed into pseudo-Cartesian coordinates are used.

Cross correlation between two consecutive images yields a motion vector field at a single instant in time, consisting of one vector for each tile. This vector field may have erroneous vectors and thus requires quality control. First, a single representative velocity is computed from the raw vector field of a single image pair. Second, that single velocity is compared against velocities from the preceding 10 minutes; and if changes in the velocity are too large, the most recent realization is eliminated.

4.1.1.4 Forecasted Shadow Map

To produce a forecast shadow map, the cloud velocity is used to advect the planar cloud map to generate a cloud position forecast for each forecast interval. The cloud position every 30seconds is then computed out to a 15 minute forecast horizon. The forecast domain is defined by a grid overlaying the plant that is 4km × 4 km in size with a resolution of 2.5 meters per forecast cell (1600×1600 cells). Each cell is resolved to a unique latitude, longitude, and altitude.

For each forecast cell, a ray is traced along a vector to the sun and the intersection with the cloud map is determined (Figure 11). If the intersected point is clear, that ground location is deemed clear, whereas if the intersection is cloudy, the ground point is deemed shaded by a cloud. Repeating the shadow mapping process for each forecast cell, a map of cloud shadows (that is, the shadow map) is constructed. The complete shadow map provides the percentage of the plant that is shaded.

Figure 11: Cartoon illustration of how the shadow map is generated from ray tracing.
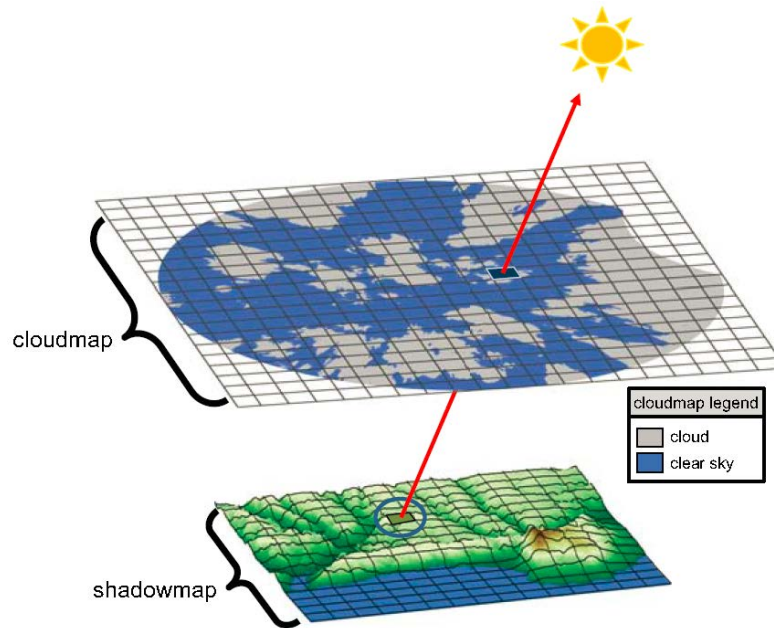
The cloud map is assumed flat. An example intersection is shown for a single grid cell in the forecast domain.

4.1.1.5 Converting the Shadow Map to Power Output

The approach assumes short-term persistence of the aerosol and cloud optical depths. Recent one second inverter power output data from the plant are used to characterize the overall atmospheric transmissivity in clear and cloudy conditions. Atmospheric transmissivity is characterized using a normalized power output given as the ratio of the inverter power output to the modeled clear sky value. This normalized power is analogous to the clear sky index (k_c) - the ratio of GHI to clear sky GHI. The plane-of-array clear sky global irradiance is determined by computing the clear sky GHI (GHI_{clr}) estimated using the Ineichen-modified Kasten clear sky model (Ineichen and Perez, 2002), breaking up the GHI into direct and diffuse following Boland et al. (2008) , and then transposing to the plane-of-array using the Muneer transposition model as described by Page (2003).

Although assuming a simple direct linear relationship between power and irradiance for use as a PV performance model is questionable, here the reverse transformation from normalized power [-] to power [MW] is directly applied and thus errors are only a function of the cell temperature difference at the current time versus the times for which the modeled clear sky power was determined.

Histograms of normalized power for two days are shown in Figure 12. Normalized power output is bimodal with a distinct clear peak, and a peak representing the modal transmissivity of the clouds. The normalized power ranges from about one tenth to above one in cases of

localized cloud enhancement. Cloud-side reflection is a common phenomenon that occurs during partly cloudy days, when the direct solar beam is unobstructed and nearby clouds reflect sunlight towards the ground. This causes observed irradiance levels to exceed those during clear skies. The cloudy and clear modes determined from the histogram over the last 2 hours are used to assign normalized power output to the shaded and un-shaded grid cells of the shadow map, respectively. The plant aggregate normalized power output $\overline{P_n}$ is computed as

$$\overline{P_n} = \frac{\sum P_n^i A_i}{\sum A_i}$$

where P_n^i is the normalized power assigned to the ith grid cell within the power plant, and A_i is the corresponding area. Normalized power [-] is computed for each cloud advection step to construct a forecast for the next several minutes. Plant power output [MW] is computed using the average normalized power $\overline{P_n}$ and adjusting the clear sky plane-of-array global irradiance for each time step.

Figure 12: Histogram of normalized power for (a) July 28, 2011 and (b) January 19, 2012 showing trimodal clear/thin cloud/thick cloud and bimodal clear/cloudy conditions, respectively.
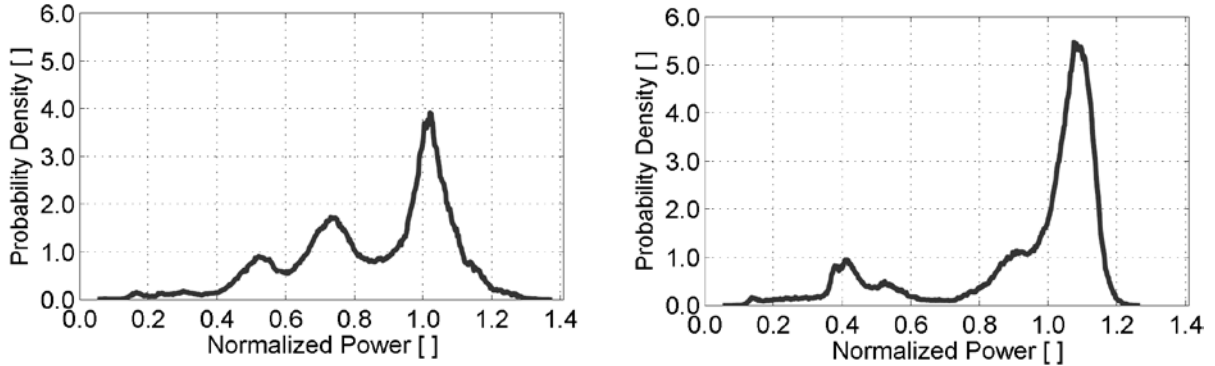
Source: AWS Truepower

4.2   Satellite Method - Cloud Motion Vector

The GOES 1 km visible imagery are used for intra-hour (1 minute resolution, 0-30 minutes ahead) and hour-ahead (30-minute resolution, 0-6 hours ahead) forecasts of site averaged-GHI at the target power plant. The conversion of the raw GOES images into GHI maps follows the method described in Perez et al. (2002). It basically consists of a two-step procedure: (1) pixel counts to cloud index conversion followed by (2) a cloud index to GHI (or clearness index) conversion. The SolarAnywhere model applies the NOHRSC and IMS snow cover data and uses the GOES infrared channels to detect and account for the presence of snow on the ground in irradiance calculations described in Perez et al. (2010). High resolution aerosol optical depth, precipitable water and ozone climatological data sets are used to calibrate the model's clear sky background for both GHI and DNI.

The short-term irradiance forecasts are produced using consecutive GOES 1 km visible satellite images to infer on the cloud motion as shown in Figure 13. Each pixel-specific cloud motion is determined using the correlation between consecutive images taken at different times (typically 30 minutes apart). Pixel-specific motion vectors are determined by minimizing the error of the difference between consecutive image-derived clearness index maps. Clearness index maps for subsequent hours, up to five hours ahead, are derived from the localized motion of clouds in individual pixels. A pixel-specific clearness index is advected in the direction of the motion determined for the considered location. The considered location is also assigned the clearness index value of a pixel located upstream from the considered location. A steady state cloud supposition is made during the forecast period during which no cloud growth or decay occurs. Future images are subsequently smoothed following the pragmatic approach described by Hammer et al. (2001). The final translation from clearness index to irradiance is then performed via the Perez et al. (2002) method.

Figure 13. Satellite cloud motion vector procedure from two consecutive images.
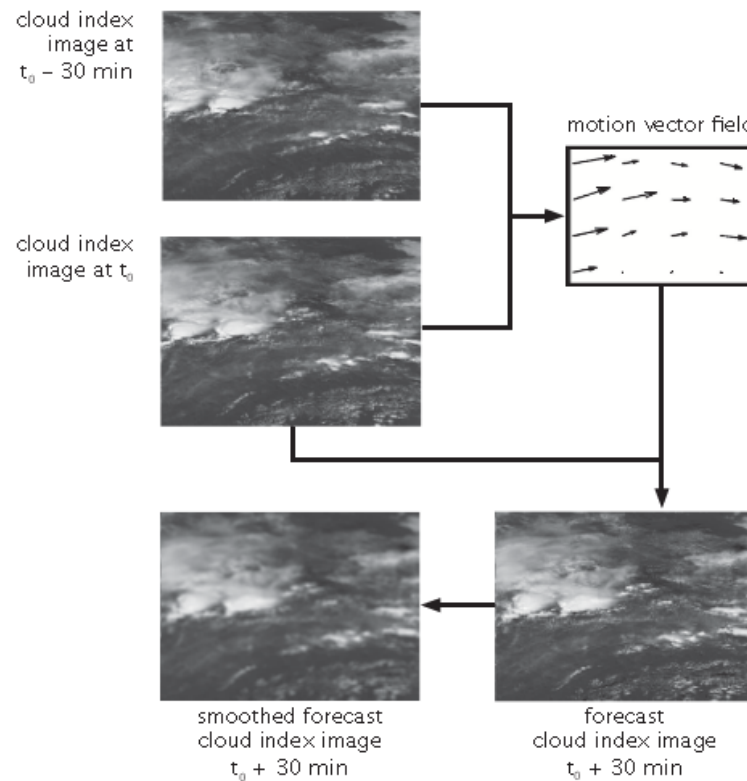
Source: AWS Truepower

The translation from irradiance to power is accomplished using PV Simulator software. Site-specific irradiance, temperature and wind speed along with PV site details (module type, inverter type, row spacing, tilt angle, DC/AC size, etc.) are all incorporated together and fed into a PV and inverter performance models. An outline of the satellite forecast derivation process is shown in flowchart form in Figure 14.

For high resolution high frequency short-term forecasts (30 minutes and less), the cloud speed determination process is identical as above. However, motion is applied at higher temporal frequency and no smoothing is applied to preserve short-term variability.

Figure 14: SolarAnywhere® satellite-based forecasting process flowchart.
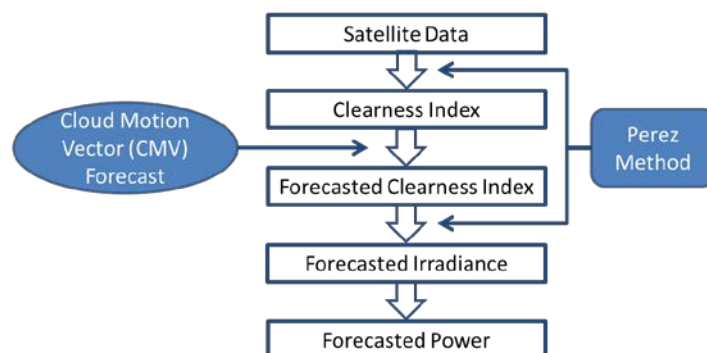
Source: Clean Power Research

4.3 Standard Cycle NWP (Regional and Global)

The NWP forecast system is composed of two major components that each consist of NWP model systems and MOS modules. Figure 15 provides a more detailed schematic of the components of the NWP subsystem. The following subsections provide an overview of the key attributes of each component of the NWP subsystem.

Figure 15: A schematic depiction of the components and data flow of the NWP system of the composite solar power forecast system.
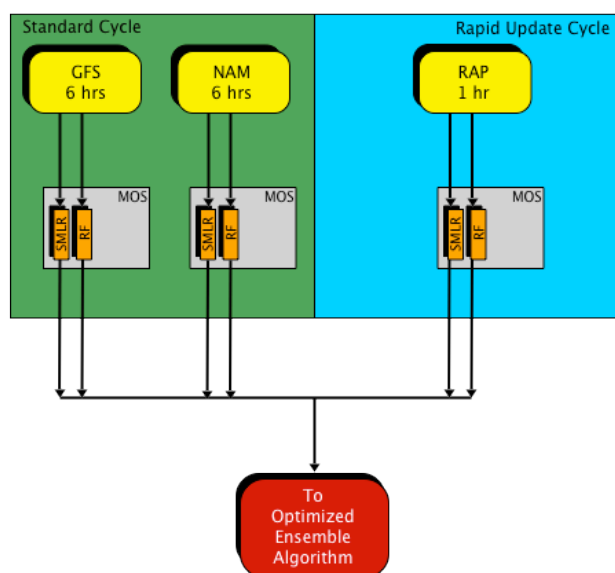
Source: AWS Truepower

The standard cycle NWP component consists of the two NWP models executed at the NOAA NCEP. These are a core component of the forecast production process used by National Weather Service. These consist of the GFS and NAM models. A summary of these two models and the output variables that were used in this project is provided.

The GFS is a global spectral model (Global Climate and Weather Modeling Branch, 2003[4]) that produces forecasts of basic atmospheric variables for all locations in the world. Since it is a spectral model, the resolution is expressed in terms of the spectral functions that serve as the basis for the solution of the model's equations. The horizontal resolution of the model changes during the course of each forecast cycle. For the 0-hour to 192-hour portion of the forecast the spectral resolution is T574, which corresponds to a physical resolution of approximately 27 km. For the 192-hour to 384-hour portion of each forecast cycle, the spectral resolution is T190, which corresponds to a physical resolution of approximately 70 km. The vertical component of the GFS grid consists of 64 unequally-spaced levels extending from the surface of the earth to 3 mb (~35 km). Approximately 15 levels are below 800 millibar (mb), or about 2 km above mean sea level, and about 24 levels are above 100 mb, or about 16 km above mean sea level).

A diverse mix of measurement data types is used in conjunction with a background state from the previous forecast cycle to create the estimation of the initial state of the atmosphere for each cycle. The measurement data is used to update the background state through a method known as Grid-Point Statistical Interpolation (DTC, 2013).

The GFS is run at NCEP four times per day. The initialization times are 0000 UTC (1600 PST), 0600 UTC (2200 PST), 1200 UTC (0400 PST) and 1800 UTC (1000 PST). A 384-hour forecast is produced by each cycle. Since the experimental forecast period in this project is 0 to 48 hours after the forecast issue time, only the first 84 hours of forecast data from each cycle are used in this project. The output from the GFS is provided at 3 hour intervals for the 0 to 192-hour forecast period and at 6 hours intervals for the forecast period that extends from 192 hour to 384 hours. Thus, it is necessary to interpolate the data if forecasts are desired for the intervening times.

Table 10 provides a listing of the variables that were extracted from GFS forecast output dataset for input into the composite forecast MOS. These parameters were extracted from the GFS output files for the location of the Copper Mountain solar generation facility for each GFS cycle whose 0-60 hour forecast period overlapped with the 24 hour period of each day in the experimental sample.

---

[4]The documentation for the most recent configuration of the model can be found at http://www.emc.ncep.noaa.gov/GFS/doc.php.

Table 10: Solar-related parameters provided in the output datasets from the NCEP GFS model that were selected as candidate predictors for MOS component of the Composite Forecast System.

Parameter ID	Description	Levels	Units
RH, ISBL	relative humidity	50 mb intervals from 1000 mb to 100 mb	%
RH, 2-HTGL	relative humidity	2-m AGL	%
RH, 0-EATM	average relative humidity	surface to top of model's atmosphere	%
CLWMR, ISBL	cloud water mixing ratio	50 mb intervals from 1000 mb to 100 mb	kg/kg
CWAT, 0-EATM	vertically integrated cloud vapor	surface to top of model's atmosphere	kg/m^2
SUNSD, 0-SFC	duration of sunshine	surface	S
PWAT, 0-EATM	vertically integrated water vapor	surface to top of model's atmosphere	kg/m^2
TCDC, 0-CCY	total cloud cover	surface to top of model's atmosphere	%

The NCEP NAM model is a grid point NWP system based on an implementation of the WRF-NMM (Janjic et al, 2010) architecture. The NAM model's domain covers all of North America. The horizontal resolution during the project period was 12 km. This atmospheric modeling system also employs the GSI method to assimilate a diverse mix of measurement data into an estimate of the state of the atmosphere at the start of each NAM forecast cycle.[5] However, the background state used to start this process is the 6-hour forecast from the previous GFS model forecast cycle.

NCEP uses the NAM model to generate 84-hour forecasts every 6 hours (i.e. 4 times per day. The forecast initialization times are 0000 UTC (1600 PST), 0600 UTC (2200 PST), 1200 UTC (0400 PST), 1800 UTC (1000 PST). NCEP provides output from the NAM model at a frequency 1 hour for the first 48 hours of the 84-hour forecast period and at 3-hour intervals for the remaining portion of the forecast period. The first 60 hours of forecast data from each cycle are used in this project. While the target forecast look-ahead period for the project is 48 hours, it should be noted that the project's forecast period is 48 hours after the forecast delivery time. Since the NAM output is not available until about 2 hours after initialization time and the NAM forecasts are updated at 6 hour intervals, there are times that the 48th hour of the composite forecast

---

[5] More information about the configuration of the NAM model can be found at
http://www.emc.ncep.noaa.gov/index.php?branch=NAM.

period requires the use of NAM forecast data that is more than 48 hours after the NAM initialization time.

Table 11 provides a listing of the variables that were extracted from the NAM forecast output dataset for input into the MOS. These parameters were extracted from the NAM output files for the location of the Copper Mountain solar generation facility for each NAM cycle whose 0-60 hour forecast period overlapped with the 24-hour period of each day in the experimental sample.

Table 11: Solar-related parameters provided in the output datasets from the NCEP NAM model that were selected as candidate predictors for MOS component of the Composite Forecast System.

Parameter ID	Description	Levels	Units
RH, X-ISBL	relative humidity	X= 25 mb intervals from 1000 mb to 50 mb	%
RH, 2-HTGL	relative humidity	2-m AGL	%
DSWRF, 0-SFC	Global Horizontal Irradiance (GHI)	surface	watts/m^2
PWAT, 0-EATM	vertically integrated water vapor	surface to top of model's atmosphere	kg/m^2
TCDC, 0-CCY	total cloud cover	surface to top of model's atmosphere	%
REFD, X- HTG	radar reflectivity	X= 1000 m and 4000 m above ground level	dB

4.4   Rapid Update NWP

The rapid update NWP was based on the NCEP RUC or RAP models. The RUC or RAP models are executed with a forecast cycle frequency that is higher than the 6-hour update frequency used by the GFS and NAM.

The NCEP rapid update model underwent a significant change during the project period although the forecast cycle update frequency of 1 hour and the horizontal resolution of 13 km remained the same through the project period. At the start of the project's one year experiential forecast period of January 1, 2012 to December 31, 2012, the NCEP rapid update model was the RUC model[6]. This model was replaced by the RAP model beginning with the 1200 UTC cycle on May 1, 2012.[7] On each cycle, the NCEP RUC or RAP model is configured to produce an 18-hour forecast. Although the NCEP rapid update model operates on a 1-hour update cycle, only every

---

[6] More information about the RUC forecast system is available at http://ruc.noaa.gov/

[7] More information about the RAP forecast system is available at http://rapidrefresh.noaa.gov.

other cycle (i.e. equivalent to a 2-hour update frequency) was used to limit the amount of processing and data storage. A schematic of the NCEP rapid update cycles actually used for each target forecast day in this project is shown in Figure 16. The output data from the NCEP rapid update model forecast cycles is available at 1-hour intervals for the entire 18-hour forecast period.

As with the other NOAA models, the RAP output dataset provided by NCEP contains a wide range of variables. The subset of variables that were selected for input into the composite forecast MOS algorithms are listed in Table 12. These include the relative humidity at a number of levels as well as the dew point depression at 2 m, radar reflectivity at two levels and the vertically integrated water vapor.

Figure 16: Time line of the forecasts from the NCEP rapid update models (RUC or RAP) relative to each days forecast target window.
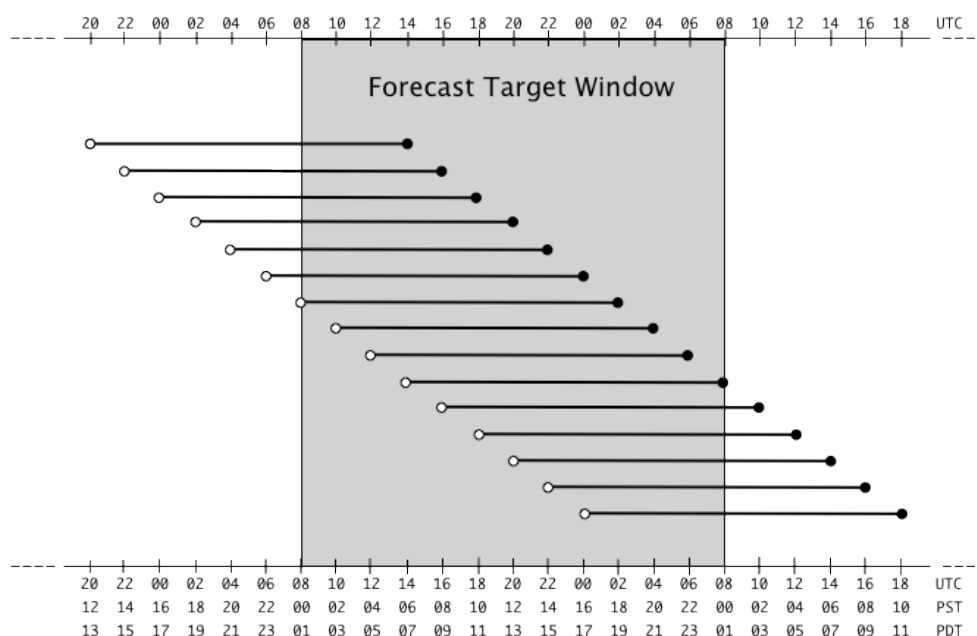
Source: AWS Truepower

Each horizontal line represents a forecast cycle. The open circles denote the initialization time for a cycle and the black circles denote the end of the forecast period for each cycle.

Table 12: Solar-related parameters provided in the output datasets from the NCEP rapid update models (RUC and RAP) that were selected as candidate predictors for the MOS algorithms in the Composite Forecast System.

Parameter ID	Description	Levels	Units
RH, X-ISBL	relative humidity	X= 25 mb intervals from 1000 mb to 100 mb	%
RH, 2-HTGL	relative humidity	2 m AGL	%
RH, 0-0DEG	relative humidity at the 0 C temperature level	height of 0°C level	%
RH, 0-HTFL	relative humidity at the highest tropospheric freezing level	height of highest tropospheric freezing level	%
DEPR, 2-HTGL	dew point depression	2 m AGL	°K
PWAT, 0-EATM	vertically integrated water vapor	surface to top of model's atmosphere	kg/m2
REFD, X- HTG	radar reflectivity	X= 1000 m and 4000 m above ground level	dB

4.5   Model Output Statistics (MOS)

MOS procedures were used to transform values of the raw output variables from each NWP model into key forecast variables: (1) GHI and (2) plant-scale solar power output. Two different MOS methods were utilized and therefore two different MOS forecasts were produced for each of the NWP models. The two MOS procedures were: (1) Screening Multiple Linear Regression (SMLR or simply LR) and Random Forests (RF).

SMLR is an approach that has been used in NWP MOS applications for many years. It was used to produce the initial set of MOS products developed by the US NWS (Glahn and Lowry, 1972). A screening multiple linear regression method was employed based on the standard approach (i.e. using a module from the R statistical software). The input into the SMLR algorithm was a subjectively selected set of predictor variables that was consisted of a combination of the raw output variables from an NWP model and variables derived from the raw output variables.

RF is based on the decision-tree regression concept (Breiman, 1974, 2001). It is a non-parametric machine learning method that improves on the decision tree approach in several ways: (1) it trains from an ensemble of trees on resampled versions of the original training set using a bootstrap technique to explore the range of dataset variability and (2) at each node of the tree, only a random subset of the data attributes are evaluated. Unlike other advanced machine learning techniques RF can simultaneously train quickly, minimize overfitting and account for

nonlinear relationships. As in the case of SMLR, the input into the RF algorithm was a subjectively selected set of predictor variables that was consisted of a combination of the raw output variables from an NWP model and variables derived from the raw output variables.

For this project, each of the respective MOS procedures is applied independently to small groups of look-ahead intervals. This enabled the MOS algorithms to adapt to changes in the underlying model's error patterns as a function of the forecast look-ahead time.

CHAPTER 5:
Composite Method Description

As noted previously two types of forecast method ensembles were employed in this project: (1) a "short-term ensemble" that was designed to support the production of 0 to 6 hour-ahead forecasts in 10-minute intervals with an update frequency of 10 minutes and (2) a day-ahead ensemble that was designed to support the production of 0 to 48-hour forecasts in 1-hour intervals with an update frequency of 1-hour. Each type of forecast is generated by constructing a composite from a set of forecasts produced by the individual members of the ensemble.

An overview of the attributes of the individual methods in each ensemble is presented in Section 5.1. This section also provides an example of the forecasts from each member of the two forecast ensembles for a selected case. The methods used to construct a composite forecast are documented in Section 5.2. Forecasts for a 1-year period extending from January 1, 2012 to December 31, 2012 were generated.

5.1   Individual Forecast Methods

5.1.1   Short Term Ensemble Method

The short-term ensemble was composed of six methods and all members produce forecasts in 10-minute intervals. However, the update frequency and the length of the look-ahead period varied substantially among the methods. This section summarizes the configuration of each method that was employed in the short-term ensemble and notes adjustments made due to the constraints.

Table 13: Forecast ensemble members for 0-6 hour ahead forecasts

Method ID	Method description	Forecast Interval (minutes)	Update Frequency (minutes)	Forecast Look-ahead Period (minutes)
TSI-CMV	Cloud motion vector technique based on Total Sky Imager data	10	10	10
Sat-CMV	Cloud motion vector technique based on satellite visible image data	10	30	10-300
RAP-LR	NOAA Rapid Refresh model with a Screening Multiple Linear Regression MOS	10	120	10-720
RAP-RF	NOAA Rapid Refresh model with a Random Forest MOS	10	120	10-720
NAM-LR	NOAA North American Mesoscale model with a Screening Multiple Linear Regression MOS	10	360	10-1080
NAM-RF	NOAA North American Mesoscale model with a Random Forest MOS	10	360	10-1080

As summarized in Table 13, the first of the six methods is labeled TSI-CMV, or simply, sky imager. The TSI-CMV method applies a cloud motion vector model to the sky image data generated by a unit operating at the Copper Mountain plant during the project period. The strength of the TSI-CMV approach is the rapid production of high spatial and temporal resolution forecasts for a limited look-ahead period. The look-ahead period is limited by the field of view of the TSI as well as the translation speed of the clouds and the rate of cloud dissipation and development. In this project, the sky imager forecasts were updated every 10 minutes and extended out to a look-ahead time of 10 minutes. That is, each forecast consisted of the GHI forecast for the next 10 minutes.

The sky imager forecasts were available only for a small portion of the 2012 period. In cases when the sky imager forecast was not available, a GHI-based clear sky index persistence forecast was used as a substitute for the TSI-CMV forecast. In addition, the length of the sky imager forecast was extended to 300 minutes (5 hours) by assuming that the clear sky index implied by the 10-minute forecast persistence through the 300-minute look-ahead period.

The second of the six methods is labeled as the Sat-CMV, or satellite. The Sat-CMV method applies a cloud motion vector model to the visible image data generated by the sensors aboard NOAA's GOES satellite. This method is conceptually similar to the TSI-CMV method but it can be effectively used for a much longer look-ahead period because of the large area covered by satellite images. Satellite forecasts generally do not perform as well as those from sky imager

method in the minutes-ahead time frame because the spatial resolution of these images is lower and may not capture small-scale cloud features.

The third and four methods are based on the application of MOS to a rapid update NWP model. The NWP model is based on NOAA's RUC/RAP model, simplified to "RAP" for the purposes of this report. This model produces a new forecast every hour, however only every other cycle was used and therefore, the update frequency was every 2 hours. The RAP produces an 18-hour forecast on each cycle. However, since the target application in this project was a 0-6 hour forecast only the output from the first 12 hours of the forecast period from every other cycle was used.

Two MOS methods were applied to the output of the RAP model. These were multivariate SMLR and RF. The SMLR method is a standard approach that has been used for NWP MOS applications for many years (Glahn and Lowery, 1972). The RF method is an advanced 'machine learning' technique that is based upon the decision-tree regression concept (Breiman, 1984, 2001). The training procedure for the application of SMLR and RF to the RAP output data was the same. Each of the 12 daily RAP initialization cycles used in this project was trained separately. Four training subsamples based on the look-ahead interval were employed for each initialization cycle: (1) 0-230 minutes, (2) 240-470 minutes, (3) 480-710 minutes and (4) 720 minutes. The data in each training sample were generated by interpolating the hourly RAP data to 10-minute intervals. The RAP data from each 10-minute interval was then assigned along with the corresponding measured 10-minute GHI data into the appropriate training subsamples based on the RAP look-ahead time. There are 24 10-minute data intervals in each of the 240-minute subsamples. This yields a potential maximum sample size of 8,760 data records for a one year training sample.

However, the training of the RAP MOS was complicated by the fact that NOAA's NCEP transitioned its rapid update NWP system during the 2012 experimental period. NCEP's previous generation rapid update model, the RUC, was the operational model until May 1, 2012, when it was replaced by RAP. Therefore, limited sample datasets were available for the statistical training after the rapid update model output changed halfway through experimental year on May 1, 2012. For model training purposes, no distinction was made between forecasts generated by the RUC and those generated by the RAP. Since no RAP data were available for years prior to 2012, 2013 data was used to train the SMLR and RF statistical models. This avoided the use of data from the forecast period in the training sample, although the 2013 data available for training was limited to January through November. These factors, including the lack of a full year of RAP training data, the exclusion of night hours, missing RAP forecast cycles from the archive for the experimental period and missing meteorological data from the target forecast site resulted in a maximum sample size of 7,196 data records for the 240-minute periods during the middle of the day and a substantially smaller sample size for those periods that encompassed the night hours.

The predictors used for the RAP-LR and RAP-RF models are listed in Table 14. The RUC/RAP dataset that was available in the archive did not include the model-generated prediction of the solar irradiance at the ground. It did have relative humidity data and several other moisture

variables. The SMLR and RF methods were directly trained with the available raw output data but the resulting forecasts had much higher error than the corresponding NAM-based forecasts whose archive included the model-generated prediction of solar irradiance. Therefore, a simple cloud transmissivity model was used to operate on the relative humidity data to produce estimates of the cloud cover in three layers, the transmissivity in those layers, and transmissivity of the entire atmosphere and the solar irradiance of the ground. These variables were used as predictors in the training of the SMLR and RF MOS models. These derived variables are denoted by the adjective 'estimated' in Table 14. Predictors used in the training of the Linear Regression (SMLR) and Random Forest (RF) MOS models for members of the hour-ahead ensemble. Three moisture variables that were directly extracted from the RUC/RAP dataset were also used as predictors: (1) precipitable water, (2) 1000-m radar reflectivity and 40-m radar reflectivity.

The fifth and sixth methods are based on the application of MOS to NCEP's NAM model. In contrast to the limited availability of the RAP historical data, the forecast data from the NAM model were available for all of 2011 and 2012. Therefore, the 2011 data were used as the training sample to develop the SMLR and RF equations for use in the production of the 2012 forecasts. As with the RAP model, the same two MOS methods (SMLR and RF) were applied to the output of the NAM model. The training procedure was the same for the SMLR and RF methods. Each of the four daily initialization cycles (0400 PST, 1000 PST, 1600 PST, and 2200 PST) was trained separately. The same four training subsamples used for the RAP training were also employed for the NAM-MOS application. The NAM output data from each cycle was interpolated from the 3-hour intervals provided in the NCEP dataset to 10-minute intervals. The 10-minute interval data was then assigned to the appropriate training subsample.

The predictors used for the NAM-LR and NAM-RF models are listed in Table 14. The first nine predictors listed were extracted directly from NAM output dataset by bi-linearly interpolating the variable from the NAM grid to the forecast location. The last three variables in the NAM section of Table 14 were added to the variables extracted from the NAM dataset.

Table 14: Predictors used in the training of the Linear Regression (SMLR) and Random Forest (RF) MOS models for members of the hour-ahead ensemble.

Predictors: NAM–LR and NAM-RF	
(1) 925 mb Relative Humidity	(7) Sfc Downward Shortwave Radiation
(2) 850 mb Relative Humidity	(8) Total Cloud Cover at Surface
(3) 700 mb Relative Humidity	(9) Model Look-ahead Hour
(4) 500 mb Relative Humidity	(10) Hour Angle
(5) 300 mb Relative Humidity	(11) Cosine (Solar Zenith Angle)
(6) 200 mb Relative Humidity	
Predictors: RAP–LR and RAP-RF	
(1) Estimated Surface Solar Irradiance	(8) 425-100 mb Estimated Transmissivity
(2) Sfc-Top Estimated Transmissivity	(9) Precipitable Water
(3) Sfc-800 mb Estimated Cloud Fraction	(10) 1000-m Radar Reflectivity
(4) 775-450 mb Estimated Cloud Fraction	(11) 40-m Radar Reflectivity
(5) 425-100 mb Estimated Cloud Fraction	(12) Model Look-ahead Minute
(6) Sfc-800 mb Estimated Transmissivity	(13) Hour Angle
(7) 775-450 mb Estimated Transmissivity	(14) Cosine (Solar Zenith Angle)

An example of the forecasts from each member of the short-term forecast ensemble is shown in Figure 17. This chart shows the forecasts of GHI produced initialized at 1000 PST (1800 UTC) on July 12, 2012 along with the measured data from the forecast site. The colored lines depict the forecasts. The black line depicts the observed 10-minute average GHI as measured by the pyranometers at the forecast target site.

The measured data suggests that the weather conditions on July 12 were characterized by a significant amount of variability in the cloudiness and solar irradiance. The 1000 PST forecast issue time was a relatively sunny period between cloud intervals. The GHI at the site had increased from an average of approximately 507 watts/m2 for the 10-minute period ending at 0950 PST to 840 watts/m2 for the 10-minute period ending at 1000 PST (the forecast initialization time). The GHI increased slightly further over the next two 10-minute intervals and then sharply decreased to a value of 173 watts/m2 by 1120 PST. This high frequency variability makes this a very difficult period to initialize a forecast model. The individual forecasts have a high degree of dispersion indicating that there is a substantial amount of forecast uncertainty.

Although it exhibits large error at the start of the forecast period, the Sat-CMV method is generally the best prediction for the midday period in this case. This is the type of case that indicates further research on forecast methods and probably better data sources are needed to improve forecasts for cases with highly variable cloud conditions.

Figure 17: Example of a set of GHI forecasts for July 12, 2012 from the short-term ensemble.
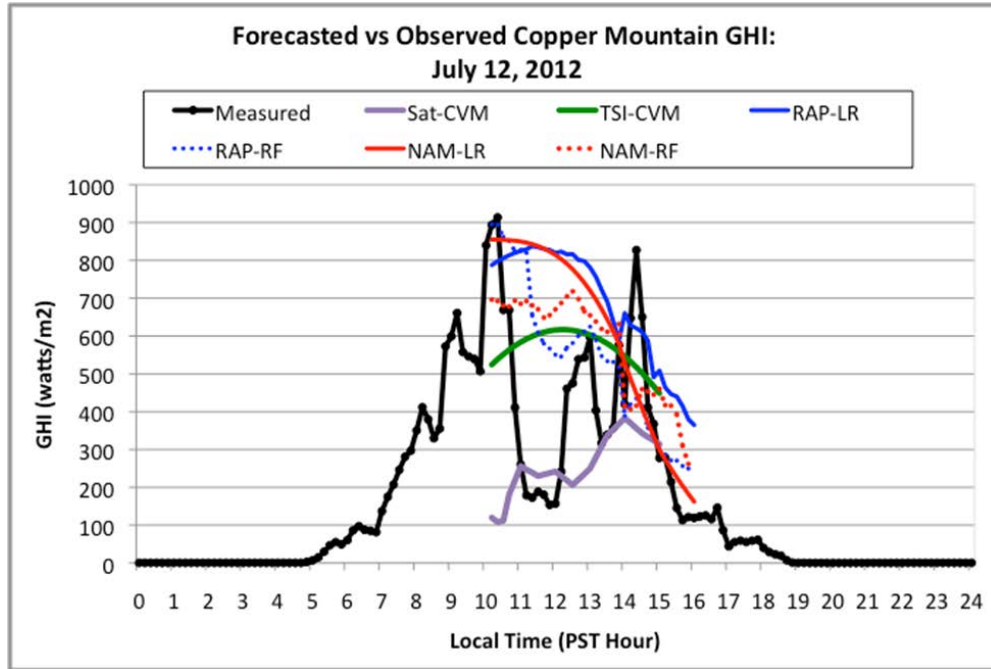
Source: AWS Truepower

5.1.2 Day-Ahead Ensemble Method

The day-ahead ensemble was composed of four methods. These methods and their salient attributes are summarized in Table 15. All produce forecasts in 1-hour intervals with an update frequency of 6 hours and a forecast length of 84 hours. Although the target length for the day-ahead forecasts was 6 to 48 hours, the forecast production procedure for the members of the day-ahead ensemble was designed to yield predictions for a 1 to 84 hour period. This allowed a forecast from each ensemble member to be available for the production of the composite forecast in cases when the data for one or more forecast cycles were not available for that member. This section summarizes the configuration of each method that was employed in the day-ahead ensemble.

49

Table 15: Forecast ensemble members for 6-48 hour-ahead forecasts

Method ID	Method description	Forecast Interval (hours)	Update Frequency (hours)	Forecast Look-ahead Period (hours)
GFS-LR	NOAA Global Forecast System model with a Screening Multiple Linear Regression MOS	1	6	1-84
GFS-RF	NOAA Global Forecast System model with a Random Forest MOS	1	6	1-84
NAM-LR	NOAA North American Mesoscale model with a Screening Multiple Linear Regression MOS	1	6	1-84
NAM-RF	NOAA North American Mesoscale model with a Random Forest MOS	1	6	1-84

The output from two NWP systems was used for the day-ahead ensemble, NAM and the GFS model. Both models are run on a 6-hour update cycle. The NAM forecast data from a particular cycle is typically available 2 to 3 hours after the initialization time and the GFS forecast data is typically available 3 to 4 hours after the initialization time. Again, two types of MOS procedures used in the short-term ensemblewere used to process the forecast data from each of these models, SMLR and RF.

The day-ahead SMLR and RF models for both of the NWP systems were trained with data from the full year of 2011. Each of the four forecast cycles (i.e. 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC) was trained independently. Since each forecast cycle begins at a fixed initialization time, this meant that each look-ahead hour for the cycle also corresponded to a fixed time-of-day hour.

Several experimental training sample configurations were tested for each of the four NWP-MOS combinations. The experiments indicated that the best performance was obtained from different training sample configurations for each NWP-MOS combination. Therefore, different configurations were employed in the ensemble. The training sample for the NAM-RF model was the entire 84-hour forecast period for each forecast cycle for the entire year of 2011. Thus, the sample size for each cycle had a maximum size of 30,665 data records (84 hours X 365 days). However, the night hours (i.e. hours with observed solar GHI of zero) were excluded from the sample and a number of forecast cycles were missing from the 2011 forecast archive. Thus, the actual training sample size was approximately 13,000 data records for each initialization cycle. The NAM-LR model was trained with a slightly different training sample configuration. In this case, the 84-hour period was divided into 28 three-hour periods and each period was trained independently. The first period was look-ahead hours 1 to 3 and the 28th period was look-

ahead hours 82 to 84. Each of these samples had a maximum size of 1,095 data records. However, the exclusion of night hours and forecast cycles missing in the archive resulted in a highly variable sample since some periods were predominantly at night. The maximum number of data records among the training periods was 472. The small sample size worked well for the SMLR approach but was more problematic for the RF approach, which generally performs much better when used with larger sample sizes.

The GFS-LR and GFS-RF models were trained with the same training sample configuration. Each initialization cycle was trained independently. The data from each cycle was divided into four subsamples: 0-23 hours, 24-47 hours, 48-71 hours and 72-84 hours. A separate model was trained for each subsample. The number of maximum number of data records for the 24-hour periods is 8,760. However, the exclusion of night hours, forecast cycles missing from the data archive and missing measurement data at the forecast site resulted in an actual sample size of about 7,000 data records for the three 24-hour training periods used for each cycle. The sample size for the one 12-hour period (look-ahead hours 72-84) had much more variability among the cycle because it was more sensitive to the initialization time of the cycle since some cycles placed this period predominately within the night hours.

The predictors from the NWP models that were used for the training of each MOS model are listed in Table 16. The predictors for the NAM model were values directly interpolated from the NAM data files for the forecast location, except the model look-ahead hour, hour angle and cosine of the solar zenith angle. However, a different approach was employed for the GFS model dataset. The GFS dataset available in the archive did not include the model-generated prediction of the solar irradiance at the ground. It did contain relative humidity data at 50 mb intervals from 1000 mb to 100 mb and several other moisture variables. The SMLR and RF methods were trained with this data but the resulting forecasts had much higher errors than the corresponding NAM-based forecasts. Therefore, a simple cloud transmissivity model was used to operate on the relative humidity data to produce estimates of the cloud cover in three layers, the transmissivity in those layers, the transmissivity of the entire atmosphere and the solar irradiance at the ground. These variables were used to train on the SMLR and RF models. These derived variables are denoted by the adjective 'estimated' in Table 16.

Table 16: Predictors used in the training of the Linear Regression (LR) and Random Forest (RF) MOS models for members of the day-ahead ensemble

Predictors: NAM–LR and NAM-RF	
(1) 925 mb Relative Humidity	(7) Sfc Downward Shortwave Radiation
(2) 850 mb Relative Humidity	(8) Total Cloud Cover at Surface
(3) 700 mb Relative Humidity	(9) Model Look-ahead Hour
(4) 500 mb Relative Humidity	(10) Hour Angle
(5) 300 mb Relative Humidity	(11) Cosine (Solar Zenith Angle)
(6) 200 mb Relative Humidity	
Predictors: GFS–LR and GFS-RF	
(1) Estimated Surface Solar Irradiance	(7) 750-450 mb Estimated Transmissivity
(2) Sfc-Top Estimated Transmissivity	(8) 400-100 mb Estimated Transmissivity
(3) Sfc-800 mb Estimated Cloud Fraction	(9) Model Look-ahead Hour
(4) 750-450 mb Estimated Cloud Fraction	(10) Hour Angle
(5) 400-100 mb Estimated Cloud Fraction	(11) Cosine (Solar Zenith Angle)
(6) Sfc-800 mb Estimated Transmissivity	

An example of a set of GHI forecasts from the members of the day-ahead ensemble is shown in Figure 18. The colored lines represent the forecasts from each ensemble member and the black line depicts the measured values. All of the forecasts are from the NWP cycle that was initialized at 0400 PST (1200 UTC) on July 11, 2012 and depict a 48-hour forecast from that cycle. The first day in this forecast period is July 11. 2012. The second day is July 12, 2012 and is the same day that was shown as an example of the forecasts from the short-term ensemble in Figure 17. The measured data for the first day in this period has a classic clear sky signature. However, as noted earlier, the measured data for the second day suggests that there was a considerable amount of cloudiness with a substantial amount of variability in the deviation of the solar irradiance from the clear sky profile during the day. All of the methods indicate that the second day would be cloudier than the first day. However, they do not predict the considerable variability in the solar irradiance over short time scales.

Figure 18: Example of a set of GHI forecasts for July 11-12, 2012 from the day-ahead ensemble.
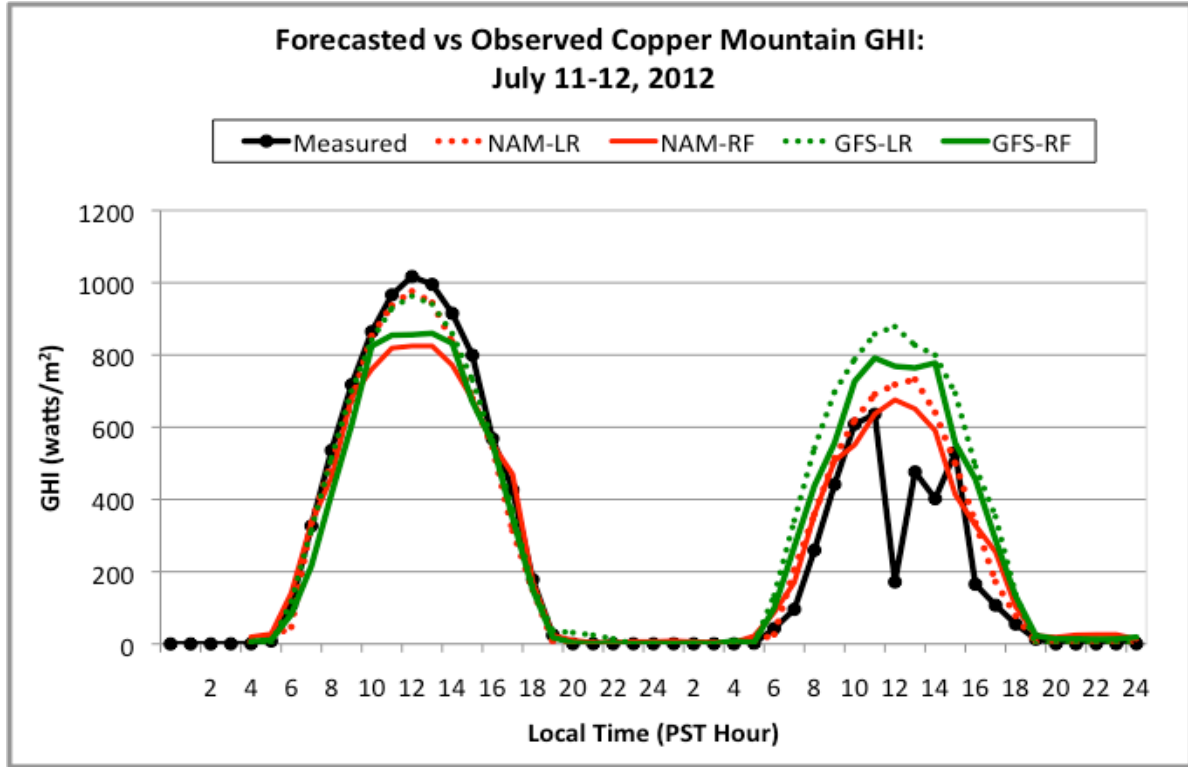
The black line is the measured GHI. The colored lines are the forecasts initialized at 0400 PST (1800 UTC) July 11, 2012 for each member of the day-ahead ensemble. The time resolution of the forecasted and observed data is 1 hour.

5.2   Composite Forecast Design

Two methods were used to construct a composite forecast from the set of forecasts produced by the members of each ensemble. One method was based on the historical error characteristics of each forecast by look-ahead time, termed inverse RMSE weighting. The other approach, using quantile regression was based on the construction of a forecast probability distribution for each forecast interval based on NWP ensemble members.

5.2.1 Inverse RMSE Weighting

The production of a composite forecast was based on the concept of weighting forecasts according to their relative performance as determined by the value of a basic error metric over a historical sample, thus an inverse RMSE weighting approach was selected. In this approach, the weight $W(F_x)$ of an individual method (x) for a particular look-ahead time is determined by:

$$W(F_x) = \frac{\dfrac{1}{(RMSE(F_x))}}{\displaystyle\sum_{n=1}^{n=N}\dfrac{1}{(RMSE(F_n))}}$$

(1)

where F_x represents the forecast to be weighted, RMSE is the Root Mean Square Error of a forecast over a representative historical period and N is the number of forecasts in the ensemble for a specific forecast cycle and look-ahead time. It is important to note that not all forecasts are available for every look-ahead period and forecast cycle. Thus, the number of forecasts included in the weighting procedure will vary by look-ahead period within a specific forecast cycle and occasionally between cycles due to missing forecasts. The composite forecast is then constructed from:

$$F_{comp} = \sum_{n=1}^{n=N} (W(F_n) * F_n)$$

(2)

Since a set of weights are computed for each look-ahead interval (e.g. for each 10-minute interval in the 0-6 hour forecast period) based on the relative RMSE values for each interval, the weighting procedure is sensitive to changes in forecast performance associated with the look-ahead time. For example, the weighting of the input methods would be expected to be substantially different for a 10-minute ahead forecast than for a 6-hour ahead forecast. A key issue is the historical period over which the RMSE is compiled. One can envision an approach where the RMSE is compiled over a rolling training sample of recent forecasts, such as the last 30 days. This could provide an indication of recent relative performance, which could be a good indicator of performance in the current weather patterns. However, there is always the risk that today's weather regime is different from the dominant regime in the recent historical sample. For the application of the method in this project the RMSE was computed over the entire experimental year of 2012. Thus, the RMSE used for a particular method and look-ahead period in Equation (1) did not vary over the year. However, the weighting of the method still varied by look-ahead time and among forecast cycle depending on the availability of the forecasts from the other methods.

5.2.2 Quantile Regression

The second method was based on a fundamentally probabilistic approach to the generation of a composite forecast. However, a deterministic forecast can also be created from the output parameters from this approach. In this approach the forecasts from each member of the ensemble are input into a quantile regression algorithm (Koenker, 1994, 2005) that develops a regression equation to predict the value of the forecast parameter (in this case either the GHI in watts/m^2 or the power production in MW) that corresponds to a specific POE. The method was configured to generate a total 13 POE levels: 98%, 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% and 2%. Many different training sample strategies can be employed to train the quantile regression. In this project, a 10-month training sample from the experimental year of 2012 was used that excluded two weeks around the target month for which the forecasts were being produced, the last 2 weeks of the prior month and the next two weeks of the subsequent month. Thus, for example, the quantile regression model used to produce forecasts for days in July of 2012 was trained on data from all days from January 1, 2012 to June 15, 2012 and August 15, 2012 to December 31, 2012. This approach was employed so that data from the days for which the forecast was to be produced were not used in the training sample.

The set of 13 POE values defines a probability distribution. The set of POE values can be transformed to a probability density function by computing incremental probability values by interpolating and differencing the probabilities associated with the POE values. This gives the probability that the forecast parameter will fall within the ranges specified by the set of POE values. These ranges of forecast parameters can then be interpolated to define a full probability density function that gives the probability that the forecast parameter will fall into any specified range.

An example of the forecasted probability distributions produced by the quantile regression algorithm is shown in Figure 19 and Figure 20. This example is for a 3-hour-ahead forecast for 1300 PST July 12, 2012 that was issued at 1000 PST July 12. The "issue time" is the same as the deterministic forecasts shown in Figure 17. The cumulative probability distribution for the GHI is depicted in Figure 19. This is essentially a direct representation (but expressed as 1-POE to be consistent with the usual format of a cumulative probability distribution) of the 13 POE values that are computed from the quantile regression. The data shown in Figure 19 can be used to derive a probability distribution that indicates the probability of the GHI falling within specified bins at the forecast valid time. The probability distribution for bins with a width of 50 watts/m2 is shown in Figure 20. This is generally a more intuitive representation of the probability data than the cumulative distribution in Figure 19. The peak probability of approximately 31% is for the 600 to 650 watts/m^2 bin. The probability distribution is also quite asymmetric with a higher mass of probabilities on the high side of the peak than on the low side of the peak. This suggests that there is a greater likelihood of thinner or scattered clouds than there is of a denser solid overcast.

Figure 19: The 3-hour ahead forecasted cumulative probability distribution for GHI for 1300 PST July 12, 2012 based on the calculation of thirteen POE values using a quantile regression algorithm with the input of GHI forecasts from the six 0-6 hour forecast methods listed in Table 13.
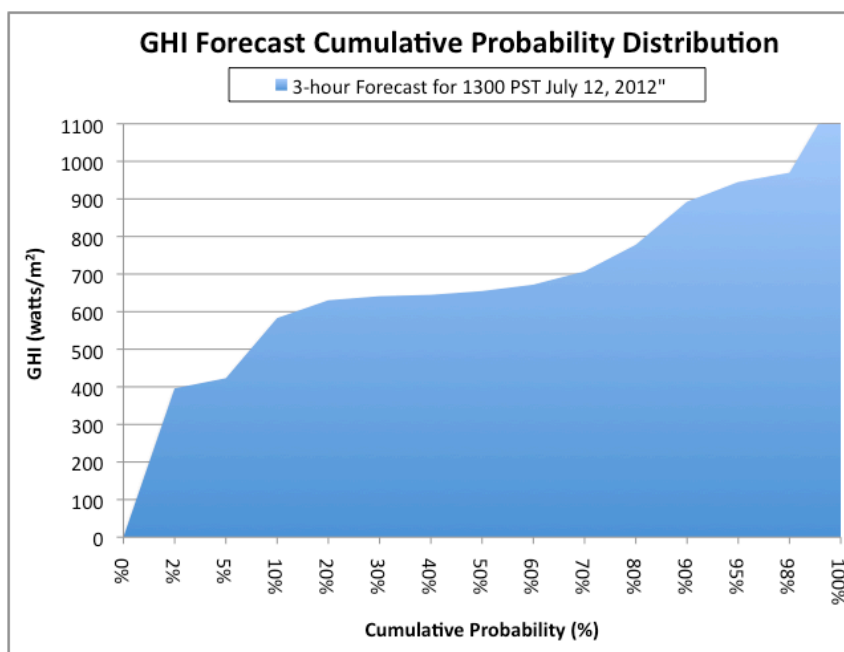
Source: AWS Truepower

Figure 20: The 3-hour ahead forecasted probability distribution for GHI for 1300 PST July 12, 2012 derived from the calculation of thirteen POE values using a quantile regression algorithm with the input of GHI forecasts from the six 0-6 hour forecast methods listed in Table 13.
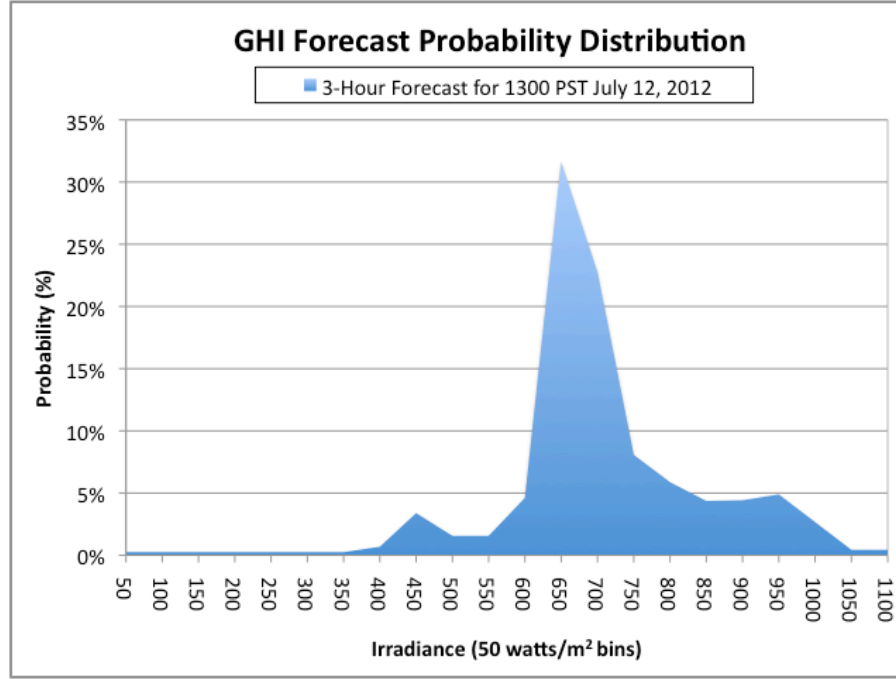
The forecast probability density function can then be evaluated using metrics that have been developed to evaluate the performance of probability forecasts. These metrics are designed to measure one or more of the key attributes of a probability forecast: reliability, resolution and sharpness. .

Instead the forecasted probability density function derived from the set of 13 POE values was used to make a deterministic forecast, which was then evaluated using standard metrics (MAE, RMSE etc.) for the evaluation of deterministic forecasts. The deterministic forecasts were generated using two different approaches.

The first approach used to select the 50% POE value as the deterministic forecast. This is the median value from the probability distribution. This would not be expected to minimize metrics such as the MAE or RMSE since it is not sensitive to asymmetries in the probability distribution.

The second approach was the computation of the "expected value". The expected value was computed from all of the probability intervals defined by the 13 POE values from

$$F_{\exp ected} = \sum_{i=1}^{N} p_i F_i$$

(3)

where $F_{expected}$ is the expected forecast value from the probability density function, N is the number of discrete probability intervals (derived from the 13 POE values) that define the

56

probability density function, F_i is the forecast value in the ith probability interval and p_i is the probability associated with that value. The training minimizes the MAE but not the RMSE. In the probability distribution example shown in Figure 20, the expected value computed from equation 3 is 673 watts/m². This is somewhat higher than the 50% POE (median) value of 655 watts/m². The expected value is higher than the median value in this case because of the previously noted asymmetry in the probability in which there is a longer and thicker tail of the distribution on the high side than on the low side.

An analogous approach could be used to compute the expected squared value of the forecast parameter. A forecasted value can be obtained by taking the square root of the expected squared value. In the probability distribution example shown in Figure 20, the expected root mean squared value computed from this procedure is 690 watts/m². This is higher than both the median and expected values due to greater weight placed upon the values in the thicker and longer tail on the high side of the distribution. This value would minimize the RMSE over a forecast evaluation sample and should be further evaluated if possible.

CHAPTER 6:
Forecast Performance Assessment

The solar irradiance and power generation forecasting challenge is closely linked to the problem of forecasting the variation of specific atmospheric variables (aerosols, clouds, winds, etc.) over short time intervals and small spatial scales. The wide variety of spatial and temporal scales that play a role in atmospheric motion, determining the variation of the key parameters within the targeted forecast volume are extremely difficult to predict.

The skill in hour-ahead forecasting is related to the prediction of small–scale atmospheric features (< 200 km in size) in the vicinity of the solar plant. Intra-hour and hour-ahead forecasting is often inferred from information about these features using meteorological and generation time series data from the target solar plant. The ability to forecast the solar irradiance and power generation over day-ahead time scales is linked to the skill of forecasting regional scale atmospheric features. These features are often referred to as synoptic and mesoscale weather systems and are the ones typically depicted in newspaper and TV weather presentations. State-of-the-art instruments and numerical models used in solar forecasting have skills that are typically limited to some range of the forecast horizon. The main goal of this project is not necessarily to identify which individual forecast method performs best within a specific time period, rather the intent was to obtain a clear understanding of the value of each forecasting method over a horizon ranging from 10 minutes to 48 hours. In this way, a composite forecast couldbe constructed that takes advantage of the strength of each method.

6.1　Error Metrics

The evaluation of solar irradiance and power forecasts involves a number of complex issues, including the determination of parameter(s) should be used as the metric(s) to measure forecast performance. The choice of metrics can have a significant impact on the characterization of forecast performance.

A wide variety of metrics are commonly used for evaluating forecast performance. Standard performance metrics include mean bias error (MBE), median error; mean absolute error (MAE), root mean square error (RMSE) and correlation coefficient (R). The equations for each metric is shown below where X is the variable of interest, e.g. GHI, power generation, the subscripts F and O correspond to the forecast and observed data, respectively.

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(X_i^F - X_i^O) \tag{1}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|X_i^F - X_i^O\right| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(X_i^F - X_i^O)^2} \tag{3}$$

$$R = \left(\frac{\frac{1}{N}\Sigma_{i=1}^{N}(X_i^F - \overline{X^F})(X_i^O - \overline{X^O})}{\sigma_{XF} \; \sigma_{XO}}\right) \tag{4}$$

where the standard deviation for any variable X is defined as:

$$\sigma = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}(X_i - \overline{X})^2} \tag{5}$$

Given the large diurnal variations in solar irradiance between sunrise/sunset and noon, it is also useful to measure the performance not only in absolute terms, but in relative term as well. In practice, solar power generation forecasts are normalized by the total capacity of the solar plant. The solar irradiance forecasts are normalized by the top of atmosphere (TOA) irradiance and effectively correspond to clearness index forecasts. The clearness index is defined as the ratio of GHI to the TOA irradiance, which is a function of time of day, i.e. sun angle. Sometimes the term extraterrestrial irradiance is used instead of TOA irradiance. The clearness index provides a measure of the opacity of the atmosphere due to clouds and aerosols. The TOA irradiance is calculated by:

$$TOA = B_0 \cdot \cos(\theta_z) \tag{6}$$

assuming a solar constant B0 at 1367 W/m2 (Luque and Hegedus 2012). Given a measured or forecasted value of GHI, the clearness index is defined as:

$$k_t = GHI/TOA. \tag{7}$$

When using the clearness index or normalized power generation instead of GHI and power generation, respectively, in the error metrics defined above, the relative mean bias error (rMBE), relative mean absolute error (rMAE) and relative root mean square error (rRMSE) are obtained.

Another distinction in selecting parameters is the sensitivity to different portions of the error frequency distribution. Some parameters are much more sensitive to outliers, i.e. forecasts with anomalously large or small errors. For example, the RMSE is quite sensitive to outliers while the median error is not. The sensitivity of the MAE parameter is between these two extremes.

In addition to the issue of different metrics providing a different picture of performance, a potential source of bias originates from the forecast systems that can be tuned to produce better

performance for a specific metric while possibly degrading the performance for other metrics. This tuning can be done by formulating a statistical technique to minimize the value of a specified optimization or cost function. Such an approach might be used to customize the forecast system to meet the needs of a specific application. However, the underlying issue is whether the evaluation metric is really linked to the user cost function. If it is, then it probably makes sense to optimize the forecast system for that metric. In this project, evaluation relied on the two most common error metrics in solar forecasting: MBE and RMSE.

6.2   Reference Forecasts

The performance of all four solar irradiance and power generation forecast approaches have been compared against reference forecasts, i.e. persistence and climatology, and evaluated against the measured irradiance and power generation data. The forecasts for the 0-30 minutes (intra-hour), or 0.5-5 hour look-ahead period (hour-ahead), are provided on a 10-minute interval with forecasts corresponding to 10-minute averages. For the 6-48 hour look-ahead period, i.e. day-ahead forecast, the forecasts are hourly-averaged values following the standard practices in forecasting.

In order to establish a baseline for the validation, reference forecasts were generated using either a persistence or a climatological model based on observed GHI and power production data from the site. A persistence forecast predicts that the current weather condition will persist and that future conditions will be the same as the present. Similarly, power production can be forecasted as a result of persistent weather conditions. Due to the strong diurnal cycle of solar irradiance, a more accurate persistence model for irradiance is one that assumes that the clearness index, ratio of GHI over top of atmosphere irradiance, rather than GHI itself will persist through time. The basic assumption is that the cloud cover, the optical depth of clouds and the aerosols will remain the same.

A climatological forecast is based solely upon the climatological statistics of a specific site or region. The method involves averaging solar irradiance or power generation statistics accumulated over many years to generate the forecast. Even though persistence forecasts can't predict ramping events, they are typically accurate for intra-hour forecast horizons while climatological forecasts provide a useful baseline for the day-ahead forecast horizon, i.e. 6-48 hour look-ahead period.

6.3   Limitations for Validation

Early in the research project, it was anticipated that each forecast method would generate forecasts for the full year of 2012, i.e. 366 days. However, the hindcast period proved challenging as the forecasting technologies were in different states of development. The sky imager method experienced limited data availability and relied on a manual process to generate forecasts due to the daily variations in the dynamical range of the TSI images and the arbitrary threshold on the red-blue ratio for cloud detection. Due to these complications, the sky imager method generated power and irradiance forecasts for a total of 60 cloudy days in 2012. On clear days, the sky imager forecast defaults to using a persistence forecast instead of the sky imager forecast, when no TSI forecast was generated.

Satellite-derived GHI forecasts were provided for the full year of 2012. However, power generation forecasts were not available due to the lack of power hindcasting capability of the SolarAnywhere forecast system at the target site. This capability was introduced later on but not in time for the validation exercise. Satellite-derived power generation forecasts were manually generated for the same 60-day sample as the sky imager method.

The NWP methods were challenged by the availability of the pre-operational NCEP RAP model outputs prior to May 2012, which limited the ability to generate the rapid update NWP forecasts. To overcome this complication, the NCEP RUC data was used to fill gaps prior to May 2012. Additionally, data available for the GFS, RUC and RAP output on NOAA's NOMADS server did not include GHI, the most important predictor for the NWP method, which potentially reduced the accuracy of the MOS. A method developed to estimate the GHI in the GFS, RUC and RAP forecasts was described in Chapter 3.3.3.

6.4   Validation of 30-day Test Period

The sky imagery and satellite forecasts were validated against persistence and climatological forecasts for the 0-30 minute forecast horizon. GHI and power generation forecasts from the TSI and satellite approaches were initially available for a total of 60 non-consecutive days in 2011-2012. The training period and validation period utilized 30 days each to provide a uniform baseline and validation analysis to compare the disparate methods.

Although clear sky conditions prevail in the Mojave Desert, the Copper Mountain plant is impacted by the North American monsoon from July to September. Thus, all forecasting methods must perform under different sky conditions to capture ramp events due to cloud shadow passage impacting the power generation. A mix of clear sky, mostly sunny, mostly cloudy and overcast conditions was selected for the initial validation of each forecasting method in the 0 to 30 minute look-ahead period. Table 17 shows the frequency of cloud cover conditions for the training and validation samples. The classification was done by calculating the daily averaged GHI as well as manually looking at the diurnal profile of GHI compared to the TOA irradiance for each day in the sample. It is likely that the non-random selection of days based on averaged sky conditions may have introduced some bias in the validation process given that the samples are likely cloudier than the site climatology. However, cloud motion vector approaches require measureable variations in the cloud optical depth in order to discern some cloud advection. Under clear sky or overcast conditions, the sky imager forecast method is not able to infer on the motion of the clouds based on two consecutive images, and therefore can not provide more value than a persistence forecast. The satellite images cover a large area (several hundred kilometers) which means there is typically some cloud structures that can be projected forward in time.

Table 17: Thirty-day training and validation sample

Category	Training Sample - Numbers of Days	Validation Sample - Numbers of Days
Sunny	10	9
Mostly Sunny	4	7
Mostly Cloudy	12	7
Cloudy	4	7

The appropriate validation time period includes only daylight hours. Nighttime values are not included in the validation. The daylight hours are defined as all times when GHI is above 0 W/m² and the sun zenith angle is within -90° to 90°. Of course, the local time of daylight hours varies significantly over the year. Therefore, some hours are night during certain parts of the year (i.e. winter) and daylight in other parts of the year (i.e. summer). The forecast interval utilized for training and validation purposes was 10 minutes.

The forecasting results from the TSI and satellite methods were generated on a 10-minute interval and compared against persistence and climatological forecasts at the Copper Mountain Solar 1 facility. For the 0-30 minute forecast horizon, the TSI and satellite forecasts were updated every 10 and 30 minutes, respectively, as shown in Figure 21.

Figure 21: Forecast update cycle for the TSI (blue) and satellite (red) methods for the 0-30 minute look-ahead period.
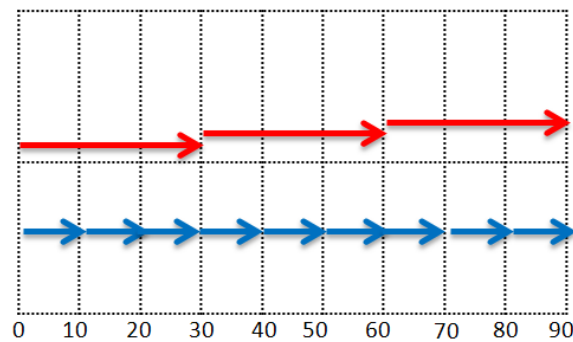
Source: AWS Truepower
The TSI and satellite forecasts are updated every 10 and 30 minutes, respectively.

The performance of the TSI and satellite GHI forecasts was binned by look-ahead time and compared to both persistence and climatological forecasts as shown in Figure 22 and Figure 23. The satellite and persistence forecasts had the lowest mean bias while climatological forecasts had the highest mean bias as shown. This also supports the hypothesis that the selected days were cloudier than the climatology. The error (RMSE) growth with lead time in Figure 23 shows

that it is difficult to improve upon a persistence forecast for the 10 to 30 minute look-ahead period. However, the RMSE of the persistence forecasts increases with lead time and becomes almost as large as the satellite forecast at a 30-minute lead time. In contrast, the satellite forecast error remains relatively stable over the 10 to 30 minute lead time. The satellite forecast doesn't degrade as rapidly as a persistence forecast based on point measurements (e.g. GHI from pyranometers or power generation from solar plant). This suggests that the resolution of the GOES satellite visible images (1 km grid cells) provides a larger scale advection/persistence of the clear sky index since it can't resolve small scale features (< 1 km). Typically, coarser spatial resolution forecasts tend to outperform high-resolution forecasts when evaluated at point measurements (Mass et al. 2002). Although a persistence forecast exhibits the best RMSE performance in the 10-30 minute look-ahead period, it will not predict a ramping event caused by cloud shadowing and therefore is of limited value when sky conditions are rapidly evolving, influenced by storm cells or frontal passages. These are typically the times that the grid system operators have the greatest need for accurate solar forecasts.

Figure 22: Mean bias error (left panel) and relative mean bias error (center panel) for GHI forecasts and power generation forecasts (right panel) by forecast horizon.
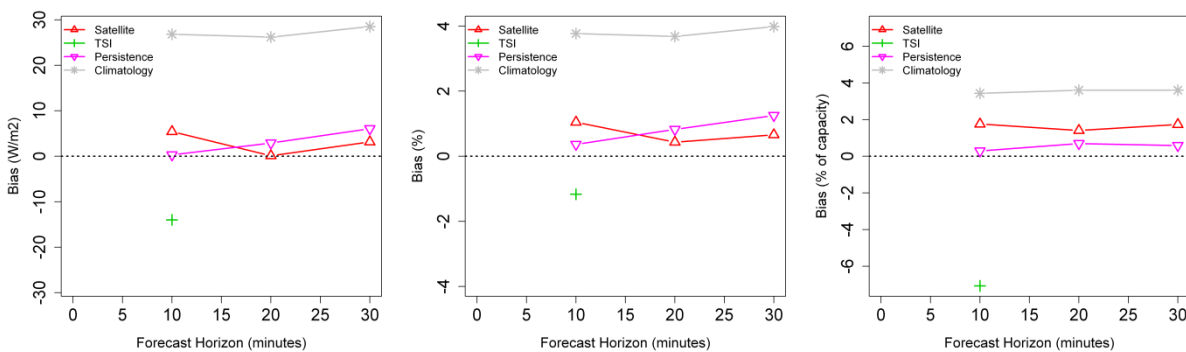
Source: AWS Truepower

All the days in Table 1 were included in the validation but the forecast methods were compared over concurrent period of time. Nighttime values are not included. The final sample size consisted of 1834 concurrent forecast records. In this study, the TSI forecasts (green + sign) were generated with a 10-minute lead time only.
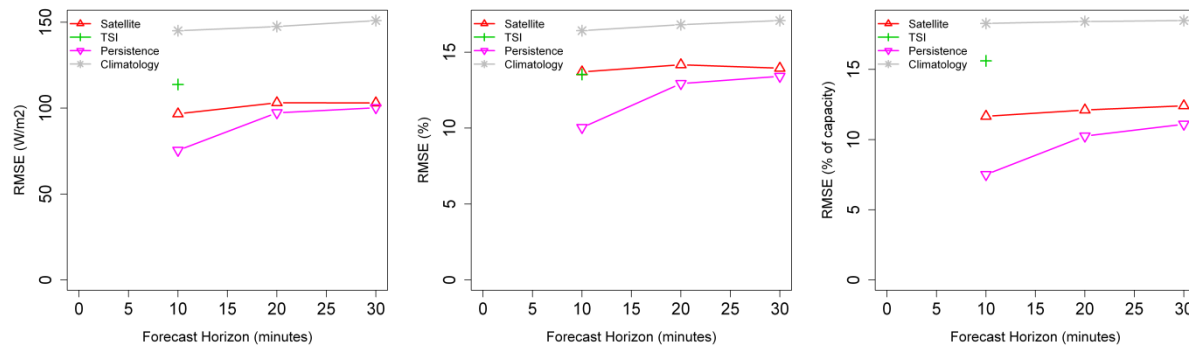
Figure 23: RMSE (left panel) and relative RMSE (center panel) for GHI forecasts and power generation forecasts (right panel) by forecast horizon.

All the days in Table 1 were included in the validation but the forecast methods were compared over concurrent period of time. Nighttime values are not included. The final sample size consisted of 1834 concurrent forecast records. In this study, the TSI forecasts (green + sign) were generated with a 10-minute lead time only.

On average, the satellite forecasts performed better than the TSI forecasts for the 10-minute look-ahead period (Figure 22 and Figure 23). In order to better understand this unexpected result, the mean bias error and RMSE of each forecast method was examined as a function of sky conditions using average sky conditions defined earlier for the training and validation sample (Figure 24). Results confirmed that the satellite approach performed better than the TSI method in all sky conditions except mostly cloudy and no large biases in the TSI forecasts for each sky condition was found. However, the climatological forecasts do show much larger mean biases than any other forecasting method for all sky conditions. As expected, the error grows with lead time (Figure 23) and the persistence forecasts do perform better than any other method, except during mostly sunny sky conditions where the satellite forecasts outperformed persistence. Although both TSI and satellite forecasts were evaluated based on the same sample of 30 non-consecutive days, it still corresponds to a relatively small sample with a mix of cloud cover conditions that are not exactly representative of the true distribution at the target plant.

Figure 24: Mean bias error (top panel) and RMSE (bottom panel) of GHI forecasts for a 10-minute lead time.
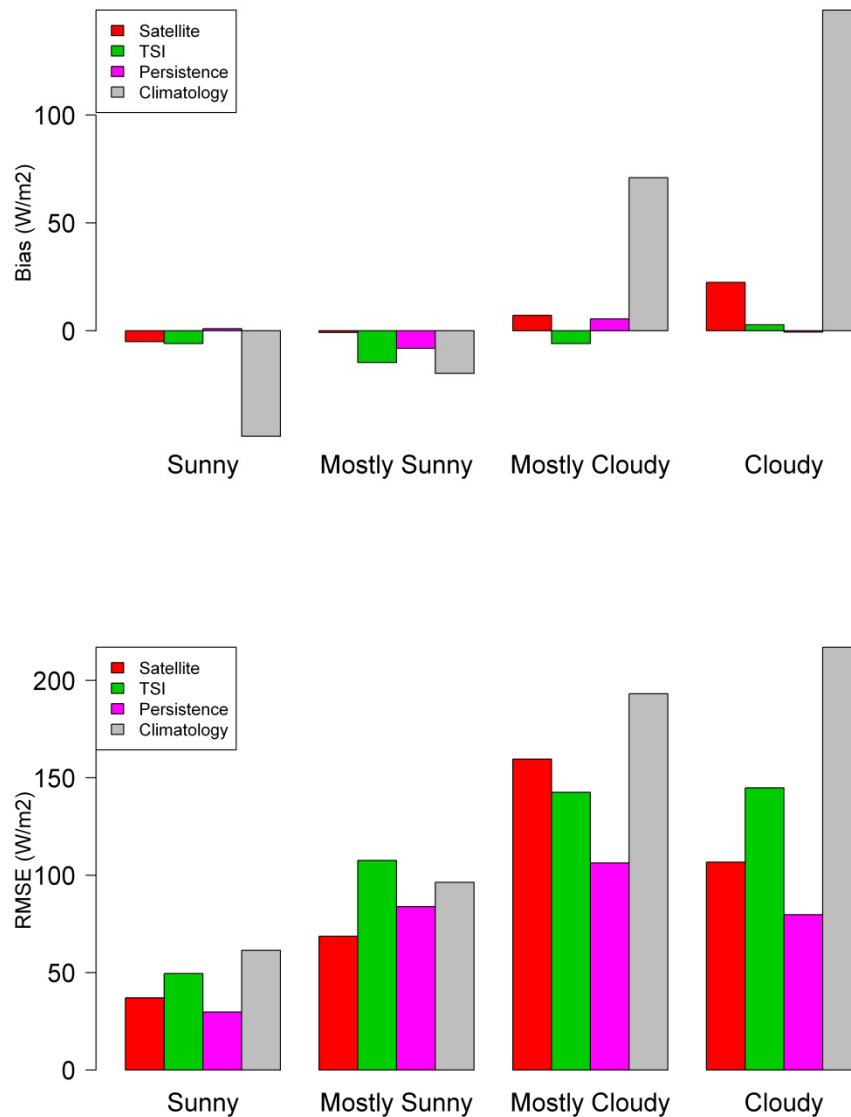
All forecast methods were compared over concurrent period of time. Nighttime values are not included.

TSI performance is negatively affected by blocked sky views due to the shadowband on the TSI mirror which blocks direct sunlight from entering the optics to improve image quality. The shadowband covers approximately 14% of the image region used for forecasting (< 80° zenith angle). Sometimes the shadowband blocks the sky region over the plant that contains clouds and as a result, minimal or no data to generate a very short-term forecast is available. The TSI camera also offers relatively low resolution and little control of the camera capture settings.

New sky imaging technology will likely remedy these issues and improve performance. Additionally, errors due to inaccurate cloud decision, cloud height, camera resolution, geometric calibration errors, cloud advection errors, and differences in cloud morphology due to viewing angle degrade performance can be improved and more research is required to further advance sky imager forecast methods. Data availability can also be improved with regular on-site maintenance to ensure the mirror is clean.

6.5   Validation of 2012 Period

Hindcasts for the full 2012 period were generated with the satellite method and NWP modeling approaches. Figure 25 shows the frequency distribution of GHI and clearness index at the target plant for all daylight hours. The daylight hours are defined as all times when GHI is above 0 W/m^2 and the sun zenith angle is within -90° to 90°. Given that GHI, rather than power forecasts were available from the satellite method for this project, the 2012 forecasts were evaluated with respect to GHI.

Figure 25: Histogram of 10-minute averaged GHI (left panel) and clearness index (right panel) based on 2012 measurements at the Copper Mountain site.
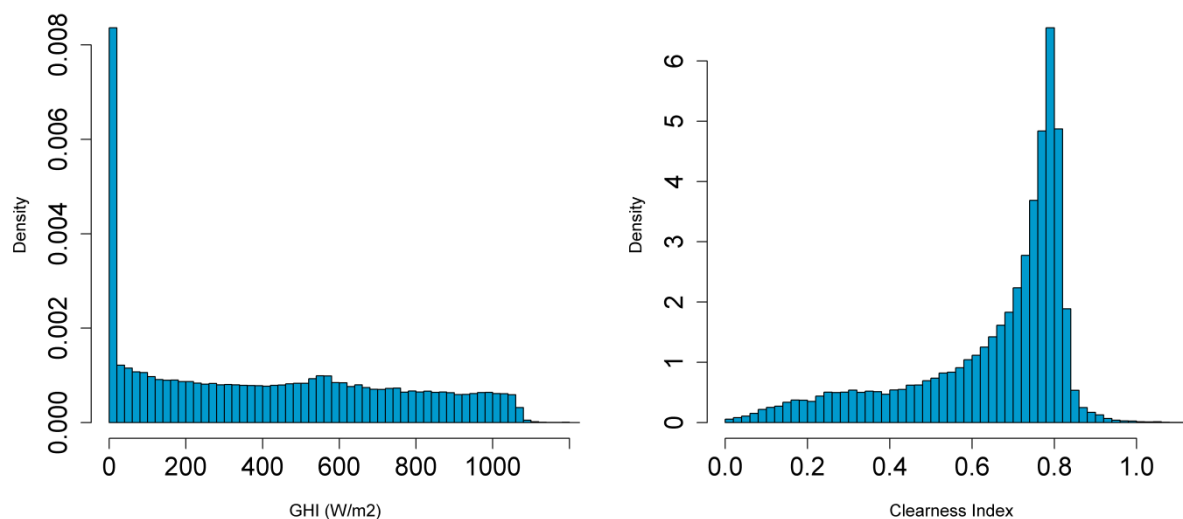
Source: AWS Truepower

All daylight hours are included.

Sky conditions are categorized into clear, partly cloudy and overcast based on the clearness index following the classification found in Lara-Fanego et al. (2012). Clear sky conditions are defined as k_t greater than 0.65, values below 0.4 as overcast conditions while in between represents partly cloudy conditions. Mathiesen et al. (2013) have defined a classification based on the clear sky index instead (GHI/GHI$_{CSK}$ where CSK means clear sky). It's roughly along the same lines as the classification of Lara-Fanego et al. (2013) if one assumes that GHI$_{CSK}$ is approximately equal to 80% of TOA irradiance. Some previous research studies found that the

diurnal GHI profiles for clear sky conditions are well approximated by the 80% of TOA irradiance profiles in eastern US (Freedman et al. 2001) and Hawaii (Beaucage et al. 2014). It is based on common Earth surface energy budget considerations where 20% of the solar irradiance is either absorbed or reflected by the atmosphere on average (Ahrens 2003).

The prevailing sky conditions at the Copper Mountain site is clear sky (Figure 26). Again, the daylight hours are defined as all times when GHI is above 0 W/m^2 and the sun zenith angle is within -90° to 90°. Since the TOA irradiance around sunrise and sunset are quite low, small errors in observed GHI or the TOA calculation can lead to large inaccuracies in clearness index. The Copper Mountain site is actually impacted by the North American monsoon originating over Mexico and spreading to the southwest US by mid-July. The monsoons are large-scale sea breeze phenomena bringing significant amount of rainfalls. Figure 27 illustrates the monthly averaged GHI peaks in June and that the summer months of July and August receive less solar irradiance than April due to the North American monsoon.

Figure 26: Frequency of clear sky, partly cloudy and overcasts conditions at the Copper Mountain site during the 2012 period. All daylight times are included.
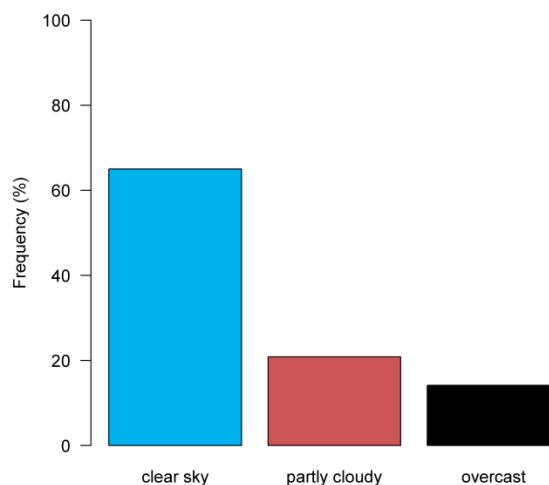
Source: AWS Truepower

Figure 27: Averaged GHI by time of day (LST) for each month at the Copper Mountain site.
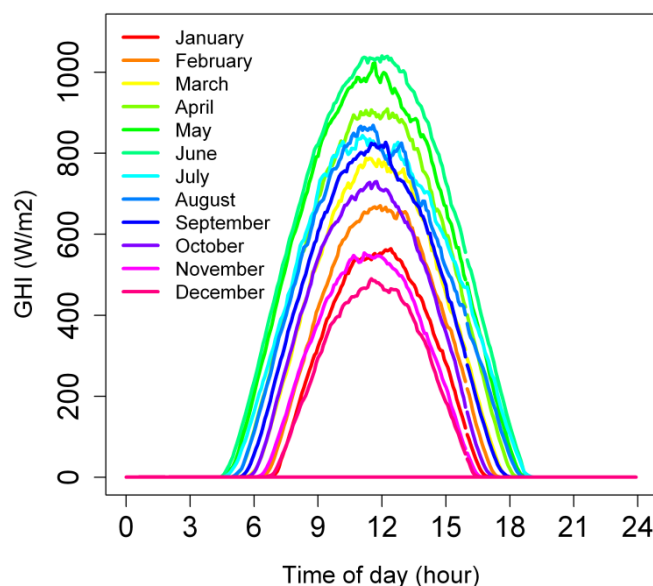
The months of July, August and September received less solar radiation than April due to the North American monsoon.

6.5.1   Validation of 30 minute to 5 hour forecasts

Satellite and NWP forecasts generated for the 30 minute to 5 hour look-ahead period were validated against both persistence and climatological forecasts. Persistence forecasts typically perform very well for minutes up to one hour lead time while climatological forecasts perform better at a 5-hour lead time or more. For the hour-ahead forecasts, the forecasts were provided as 10-minute averages, not instantaneous values. The forecast approaches were available for the entire year of 2012. The NWP-RAP forecasts were updated every 2 hours at specific times (from 0 to 22 UTC), unlike the satellite forecasts which was updated every 30 minute. Hence, the forecasts results were evaluated using the concurrent forecast delivery and horizon times of the 2-hour update cycle, as it was the lowest common denominator among all forecasting approaches.

The validation included only daylight hours, from approximately 6am to 8pm local time. Daylight hours are defined as all times when GHI is above 0 W/m² and the sun zenith angle is within -90° to 90°. In the winter, the day length is approximately 10 hours while in the summer there are approximately 15 hours of daylight at the target site. A first look at the diurnal cycle of forecasted GHI compared to observed GHI shows that forecasting methods provide reasonable diurnal profiles (Figure 28). The satellite, NWP-RAP and climatology forecasts have similar profiles to the observed curve but the combination of low sample size in the early morning hours. Less frequent updates of the RAP model as opposed to the satellite-based method produces a "kink" in the profile around 8:30am LST likely due to the two-hour RAP updates used in this study. The satellite-based forecasts are updated every 30 minutes.

Figure 28: Averaged GHI by time of day in local standard time (LST). Forecast methods were compared over concurrent period of time only.
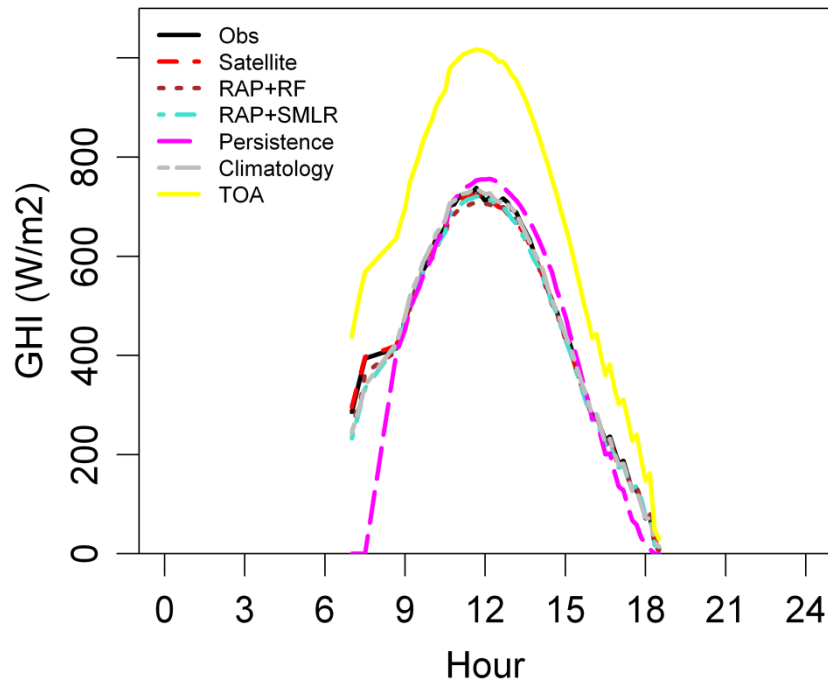
Nighttime values are not included. The final sample size consisted of 27116 concurrent forecast records for the entire hour-ahead forecast horizon.

Solar radiation reaching the ground varies by seasons as the day length and changing sky conditions as a result of synoptic patterns. The error patterns from month-to-month and hour-to-hour are similar for the satellite, RAP and climatological forecasts as shown in

Figure 29. However, the persistence forecasts exhibit much higher error than the other forecasting methods. All forecast approaches perform well in May and June when clear sky conditions are prevailing. In July and August, errors are much larger than for any other months (see Figure 29, left panels) even if the GHI forecast errors are normalized by the TOA irradiance to remove some seasonal effects (see Figure 29, right panels). The satellite, RAP and climatological forecasts still perform worse in July and August than the rest of the year. Of course, partly cloudy conditions during the North American monsoon are more challenging to forecast than the prevailing clear sky conditions of May and June.

GHI forecast error is typically larger around midday or noon as an artifact of the clearness index calculation and because GHI peaks around noon. Large errors occur in clearness index forecasts around sunrise and sunset due to low TOA irradiance value around sunrise and sunset, where small errors in GHI forecasts can appear as large errors in k_t when normalizing by low TOA irradiance values (see equation 7). Overall, forecasting error around sunrise and sunset are of

minimal impact on the grid as the power generation at the solar plant is small given the low irradiance, limits on panel tilts and shading effects.

Figure 29: RMSE (left panels) and relative RMSE (right panels) by month and hour for GHI forecasts.
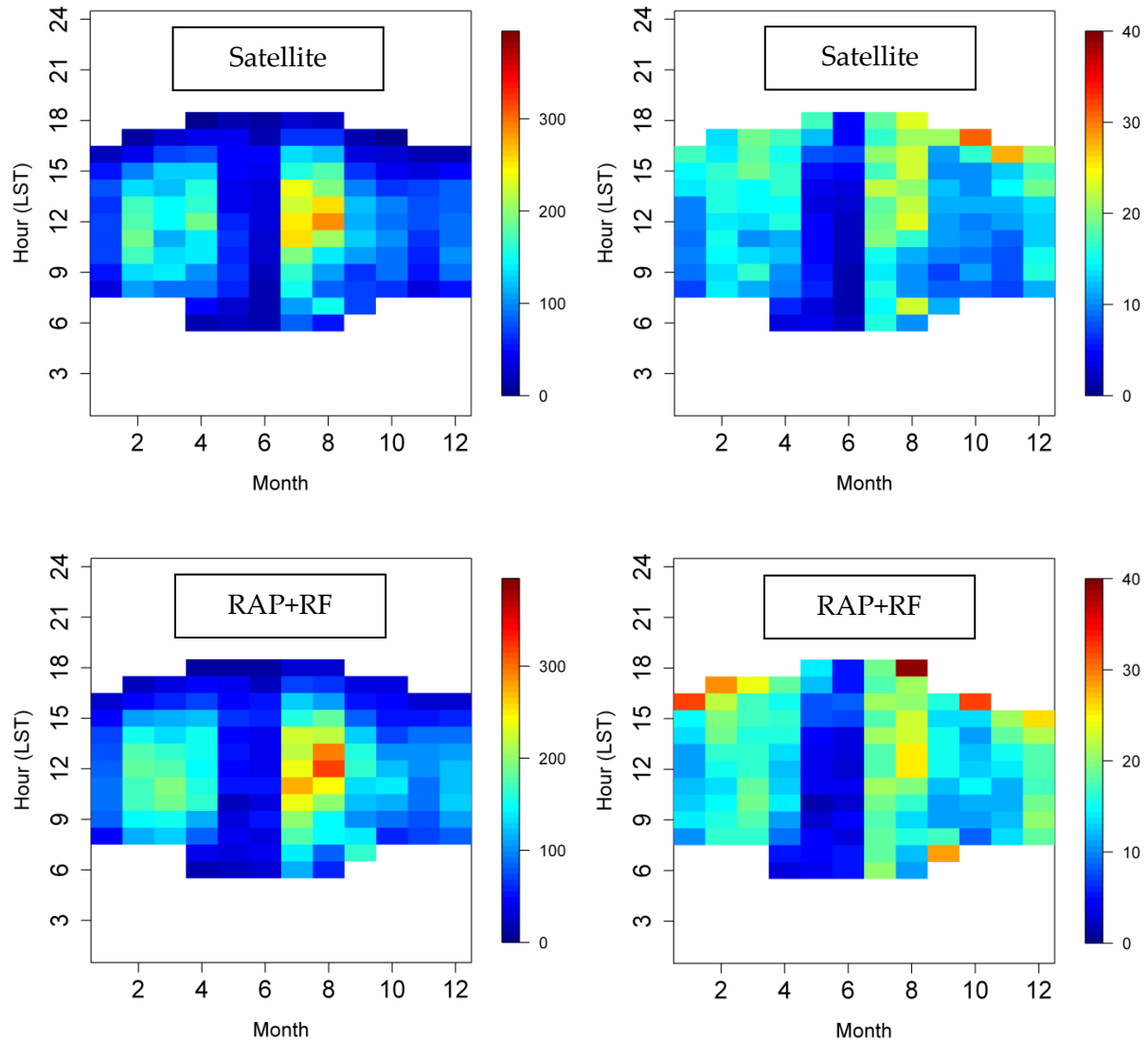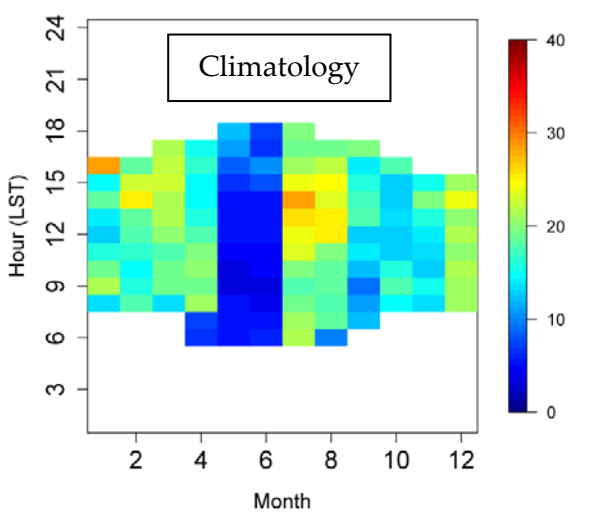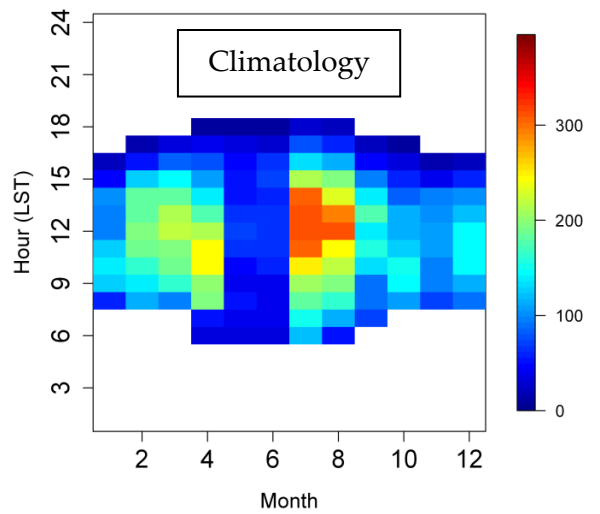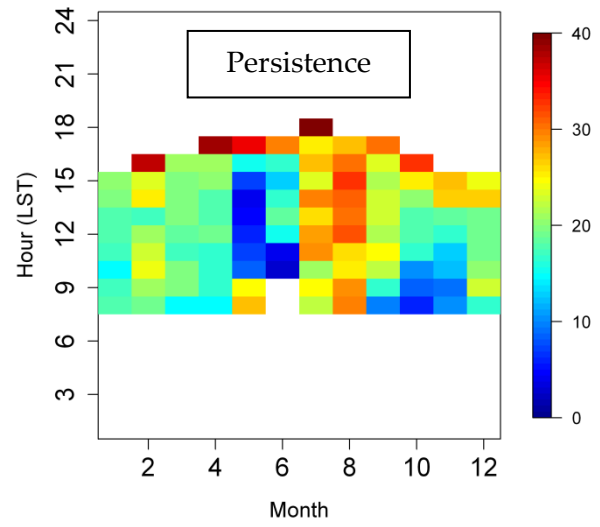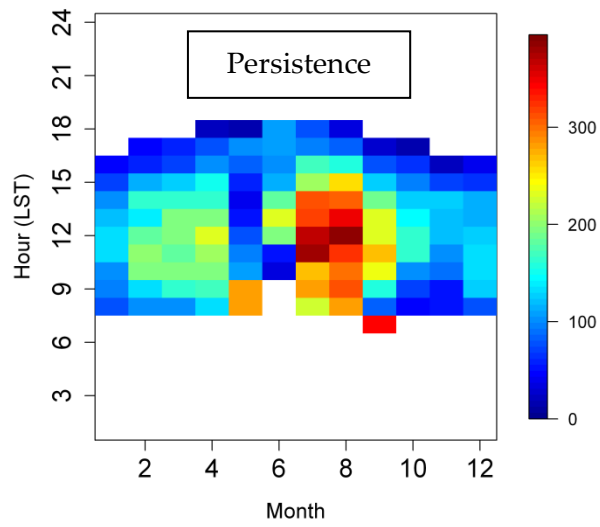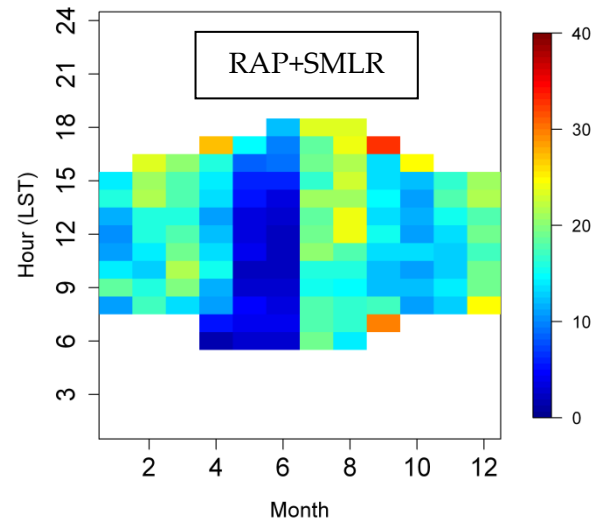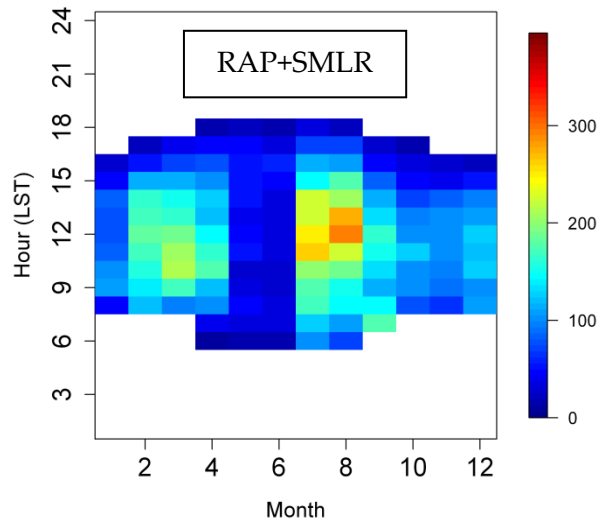
For Figure 29 the units are in W/m2 for GHI and % for clearness index. For visualization purposes, the color scale for the clearness index (right panels) was limited to 40% although it can reach values well above 40% around sunrise and sunset. From top to bottom, the pair of GHI and clearness index RMSE corresponds to the satellite, RAP+RF, RAP+SMLR, persistence and climatology forecasts, respectively.

Following standard procedures in solar forecasting, the GHI (or k_t) forecasts were evaluated in terms of their mean bias and RMSE (Perez et al. 2010, Lara-Fanego et al. 2012). The growth in error of the mean bias and RMSE for the 30 minute to 5 hour forecast horizon is shown in Figure 30 and Figure 31, respectively. Figure 30 indicates that the persistence forecast loses value over the climatological forecast after a 2-hour look-ahead time. Overall, all forecasting methods have low relative mean bias (Figure 30, right panel). The relative mean bias of the climatological forecasts is close to 0% as would be expected given that the validation period covers a full year.

The satellite forecasts show lower RMSE values than the RAP forecasts up to 3 to 4 hours of lead time. The RMSE growth with look-ahead time is relatively stable for the RAP forecasts as well as the climatological forecasts within the 30 minute to 5 hour forecast horizon. Interestingly, the satellite RMSE tends to increase rather linearly with forecast horizon. A potential explanation for these two different patterns in error growth is that the dynamic nature of cloud formation and dissipation is not taken into account in the satellite cloud motion vector forecast algorithm. The RAP forecasts initially have higher RMSE values than the satellite forecasts at a 30-minute lead time. However, the RAP models include several state-of-the-art physics parameterization schemes such as solar radiation, land surface models, PBL turbulence, cloud convection and microphysics that allow the models to simulate the evolution of clouds under different atmospheric conditions. It would be an over-statement to claim that physics schemes can capture accurately the development of clouds but modern NWP models are continually improved in the scientific community. In any case, the satellite and RAP GHI forecasts perform better than the persistence and climatological forecasts (Figure 31, left panel). A slightly different conclusion is reached when looking at the relative RMSE (Figure 31, right panel) given that the RAP forecast accuracy is more on par with the climatological forecast accuracy. The relative RMSE by look-ahead time show oscillations with a period of 30 minutes. However, this pattern is not seen in the RMSE plot. The oscillations in the relative RMSE plot disappear when filtering for forecasts between 9am and 4pm LST. The GHI forecast errors near sunrise and sunset, correspond to small TOA irradiance values, resulting in a large relative RMSE.

Figure 30: Mean bias (left panels) and relative mean bias (right panels) of GHI forecasts by look-ahead period between 30 minutes to 5 hours.
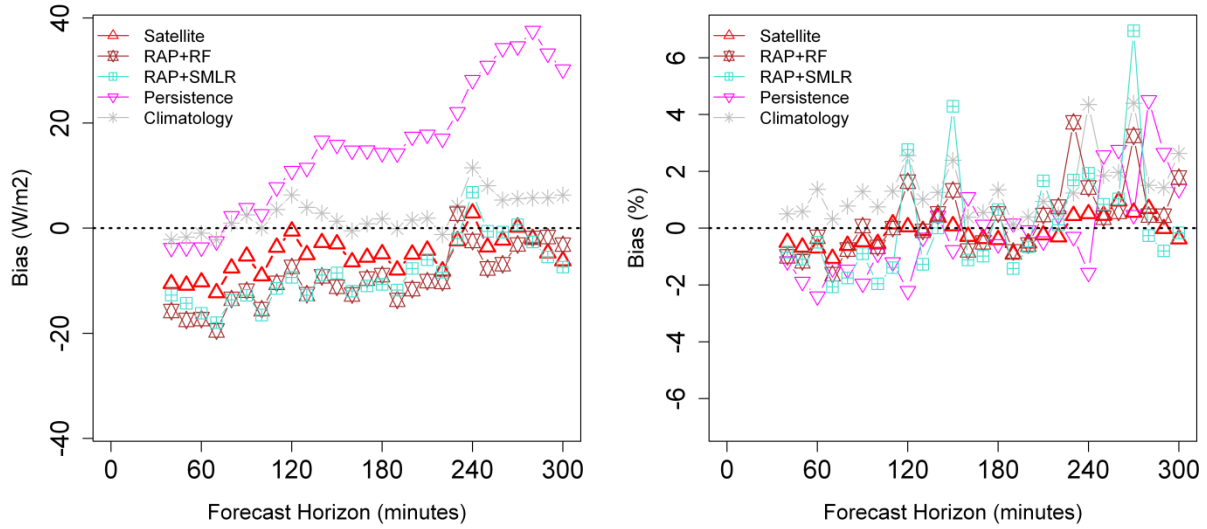
For visualization purposes, we have limited the range of bias values to provide a better comparison between the satellite, RUC/RAP and climatological forecasts. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

Figure 31: RMSE (left panels) and relative RMSE (right panels) by look-ahead period between 30 minutes to 5 hours.
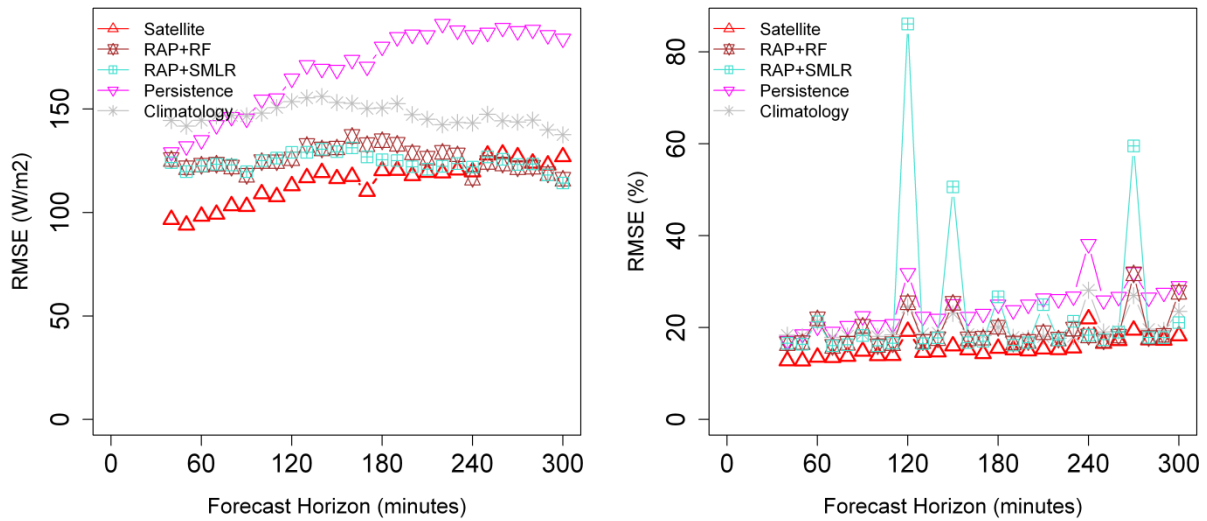
For visualization purposes, we have limited the range of RMSE values to provide a better comparison between the satellite, RUC/RAP and climatological forecasts. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

73

The forecast errors for the entire 30-minute to 5-hour look-ahead period can be summarized in a Taylor diagram (Taylor 2001). The Taylor diagram includes the RMSE, correlation coefficient against onsite measurements as well as the standard deviation of the forecast variable (e.g. GHI) for each dataset, into a single plot. Simply put, the best forecasts will lie as close as possible to observed data point located on the X axis. For the 30 minutes to 5 hour look-ahead period, the satellite forecasts show the best performance overall while the RAP model forecasts follow closely (RAP with derived GHI field) (Figure 32). The satellite and RAP forecasts surpass in accuracy the climatological forecasts. According to the Taylor diagram, it is not possible distinguish the accuracy of the RF or SMLR MOS versions of the RAP forecasts.

Figure 32: Taylor diagram of 10-minute average GHI forecasts for the 30 minutes to 5 hours forecast horizon.
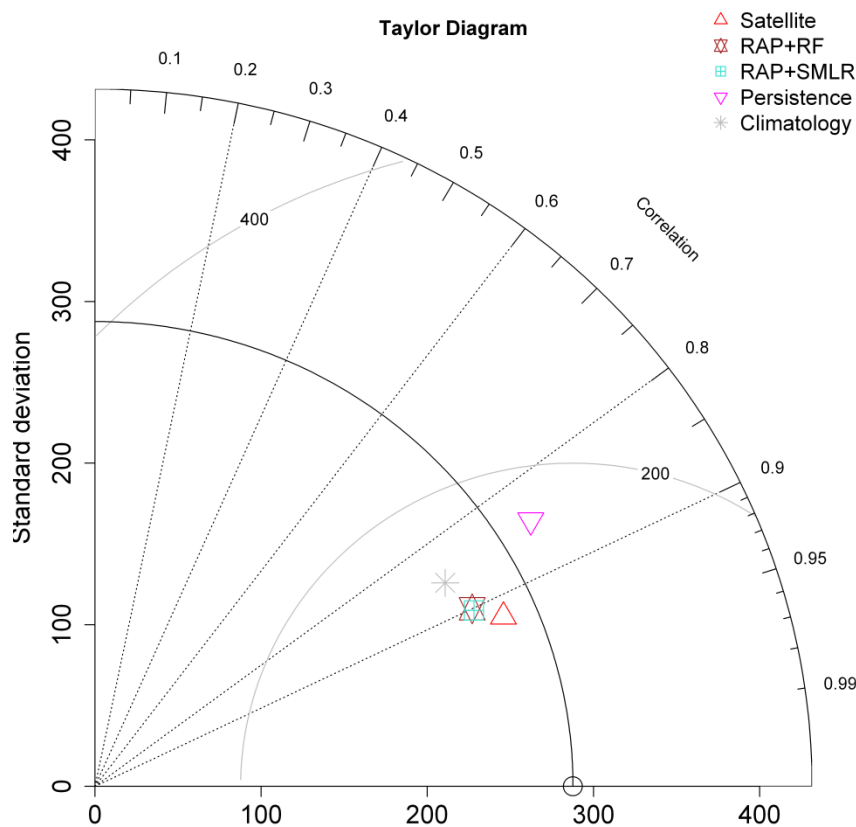
Source: AWS Truepower

The angle corresponds to the correlation coefficient (0 to 1), the radii the standard deviation (W/m2) within each forecast dataset while the light grey arcs refer to the RMSE (W/m2). The best forecasts will lie as close as possible to the observed data point (black circle). All forecast methods were compared over concurrent period of time. Nighttime values are not included.

6.5.2 Validation of 6 to 48 hour forecasts

NWP-based forecasts for the 6 to 48 hour look-ahead period were validated against a 24-hour persistence forecast and a climatological forecast. A climatological forecast of solar irradiance tends to outperform a clearness index-based persistence forecast after 1 to 2 hours of lead time. A 24-hour persistence forecast was used as a reference for the 6-48 hour forecast horizon. The persistence forecast used the previous day's profile as the forecast for the next day. Thus, the forecast for each interval of the next day was the value for the corresponding interval on the previous day. For instance, the persistence forecast for 11 AM to noon tomorrow is the 11 AM to noon value observed today.

Following standard practices in solar (and wind) energy, the forecasts for the 6-48 hour time horizon correspond to hourly averages, not instantaneous values. In addition, the "period ending" naming convention for the hourly averages has been used. Thus, the 1400 LST forecasts are the hourly average for the 1300-1400 LST period. The difference between period-beginning and period-ending hourly averaged solar irradiance (and power generation) forecasts can be substantial given the diurnal changes in sun angle, i.e. rotation of the earth. Therefore, the observed GHI, power generation data and forecasts were processed to hourly averages - period ending.

The NWP model forecasts were available for the whole year of 2012. The validation included only daylight hours, when GHI is above 0 W/m² and the sun zenith angle is within -90° to 90°. The diurnal profiles as shown in Figure 33. All forecasting methods illustrate realistic diurnal profile that aligns with the observed; however the NAM+RF results appear to underestimate the solar resource around midday.

Figure 33: Averaged GHI by time of day in local standard time (LST).
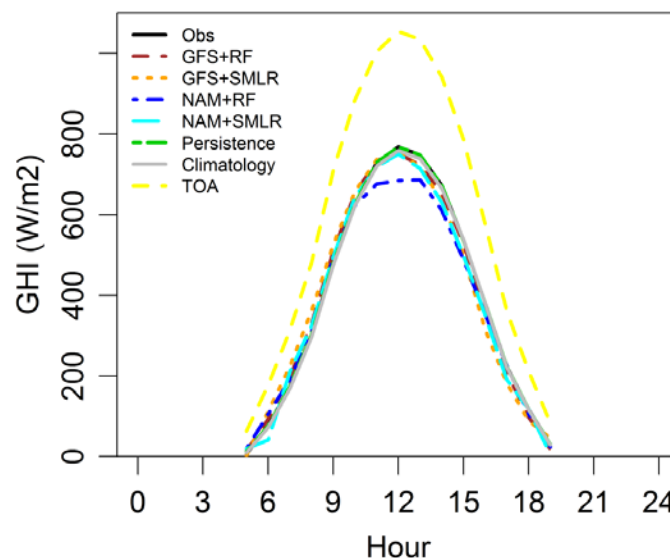
Source: AWS Truepower

Forecast methods were compared over a concurrent period of time. Nighttime values are not included.

The error by month and time-of-day are similar for all NWP-based forecasts, GFS and NAM, regardless of the MOS applied. As expected, the GHI forecast errors are typically larger around midday and seasonally higher in July and August (see Figure 34, left panels). When the GHI forecast errors are normalized by the TOA irradiance, the midday NWP-based and reference forecasts perform as well as other time of day, excluding sunrise and sunset (see Figure 34, right panels). The 24-hour persistence forecasts perform very well in May and June, but substantially worse than NWP and climatology forecasts in July and August. This strongly suggests that the partly cloudy conditions during the North American monsoon season are more challenging to forecast than the prevailing clear sky conditions during May and June. As expected, some large errors in clearness index forecasts occur in the hours near sunrise and sunset due to low TOA irradiance values which cause small errors in GHI forecasts to appear as large errors in kt (see equation 7).

Figure 34: RMSE (left panels) and relative RMSE (right panels) by month and hour for GHI forecasts.
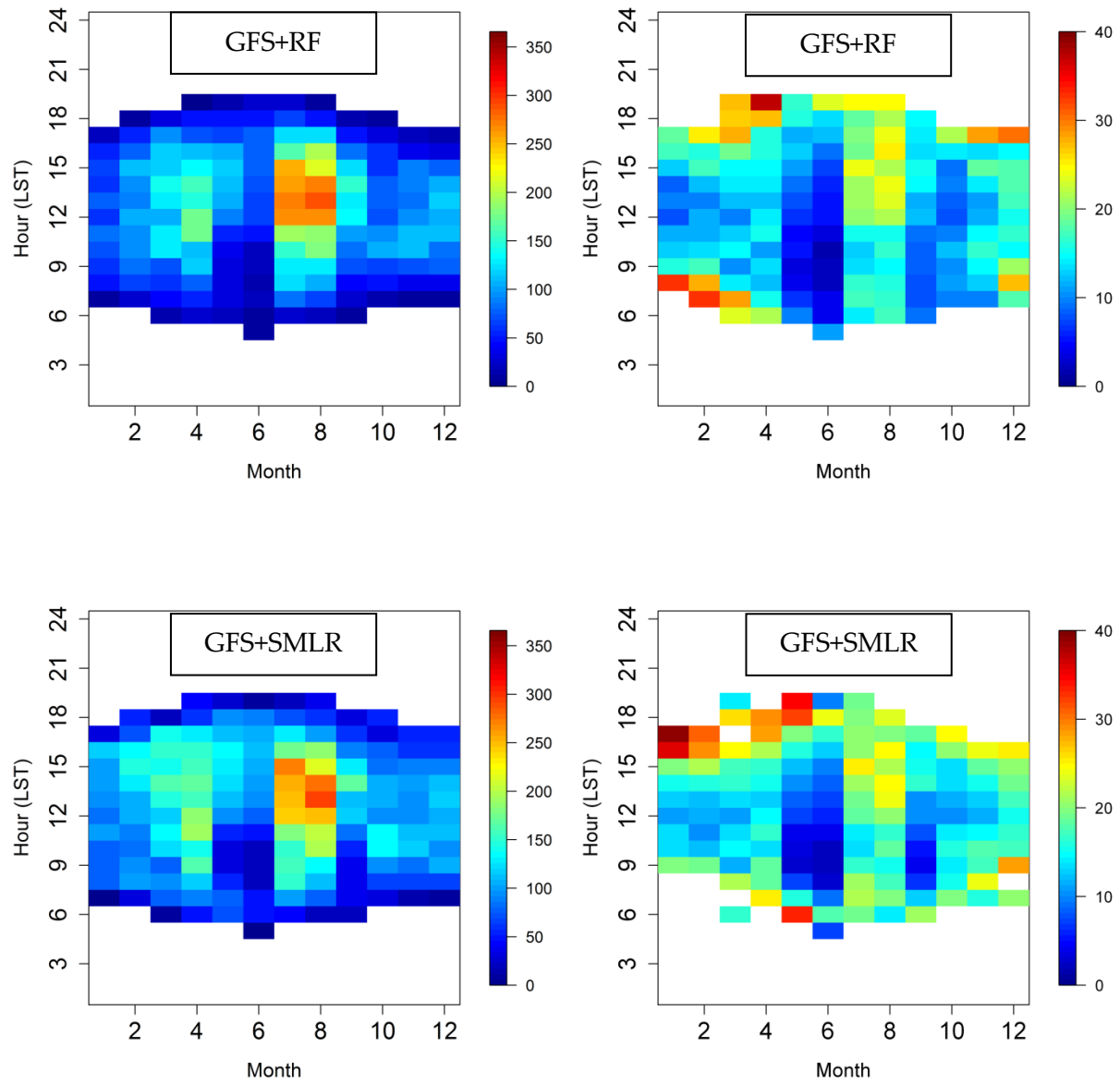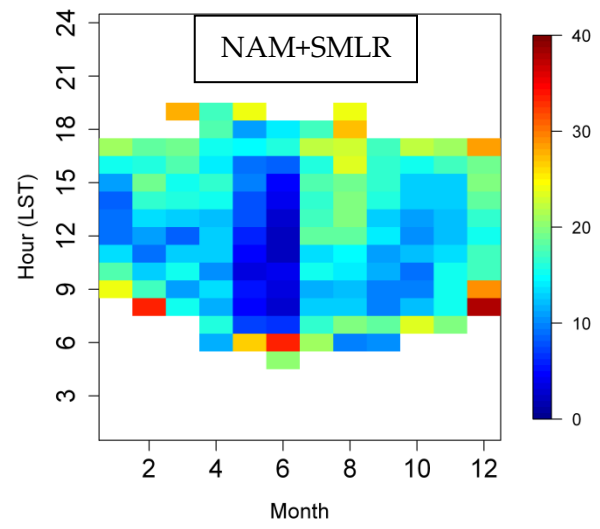
Source: AWS Truepower

The units are in W/m2 for GHI and % for clearness index. For visualization purposes, the color scale for the clearness index (right panels) was limited to 40% although it can reach values well above 40% around sunrise and sunset. From top to bottom, the pair of GHI and clearness index RMSE correspond to the GFS+RF, GFS+SMLR, NAM+RF, NAM+SMLR and climatology forecasts, respectively.

Another important characteristic of solar forecasts is their error growth with look-ahead time. Forecasting errors such as MAE or RMSE tend to grow with lead times as shown in Perez et al. (2010) and Lara-Fanego et al. (2012). Following standard procedures in solar forecasting, the GHI (or k_t) forecasts were evaluated in terms of their mean bias error and RMSE. Results indicate oscillations in the mean bias of the hourly GHI forecasts are due to a combination of

79

two factors: nighttime values were removed from the analysis and the NWP forecasts update cycles. NWP provide forecasts 4 times a day at specific hours, i.e. 0, 6, 12 and 18 UTC which correspond to 16, 22, 4 and 10 LST, respectively. In order to remove the oscillations in error statistics by lead time, the hourly averaged forecasts were averaged into 6-hour blocks. The 6-hour averaged error statistics provides a much clearer picture of the error growth with lead time for each forecasting method. Figure 35 shows the mean bias error and relative mean bias error for all the NWP-based and climatological forecasts for the 6 to 48 hour forecast horizon. The biases show considerable variations by look-ahead time. The GFS with RF has the smallest bias while NAM with RF has the largest bias. The mean bias of the climatological forecasts should remain at 0 but given some missing values, the sample of valid observed and forecasted GHI values available for validation do not represent the exact same mix of sky conditions as the whole 2012 period. Nevertheless, the mean bias of the climatological forecast remains low with a relative mean bias below 1%.

Figure 35: Mean bias error (left panels) and relative mean bias error (right panels) by forecast horizon for GHI forecasts averaged over a 6-hour block.
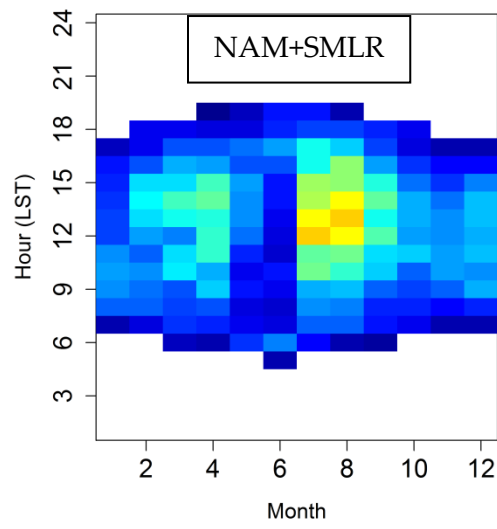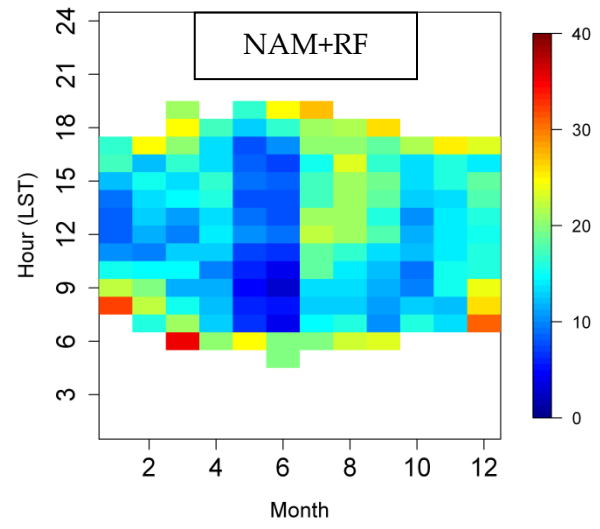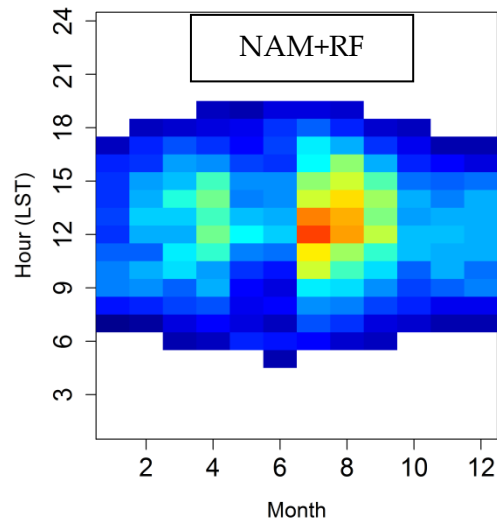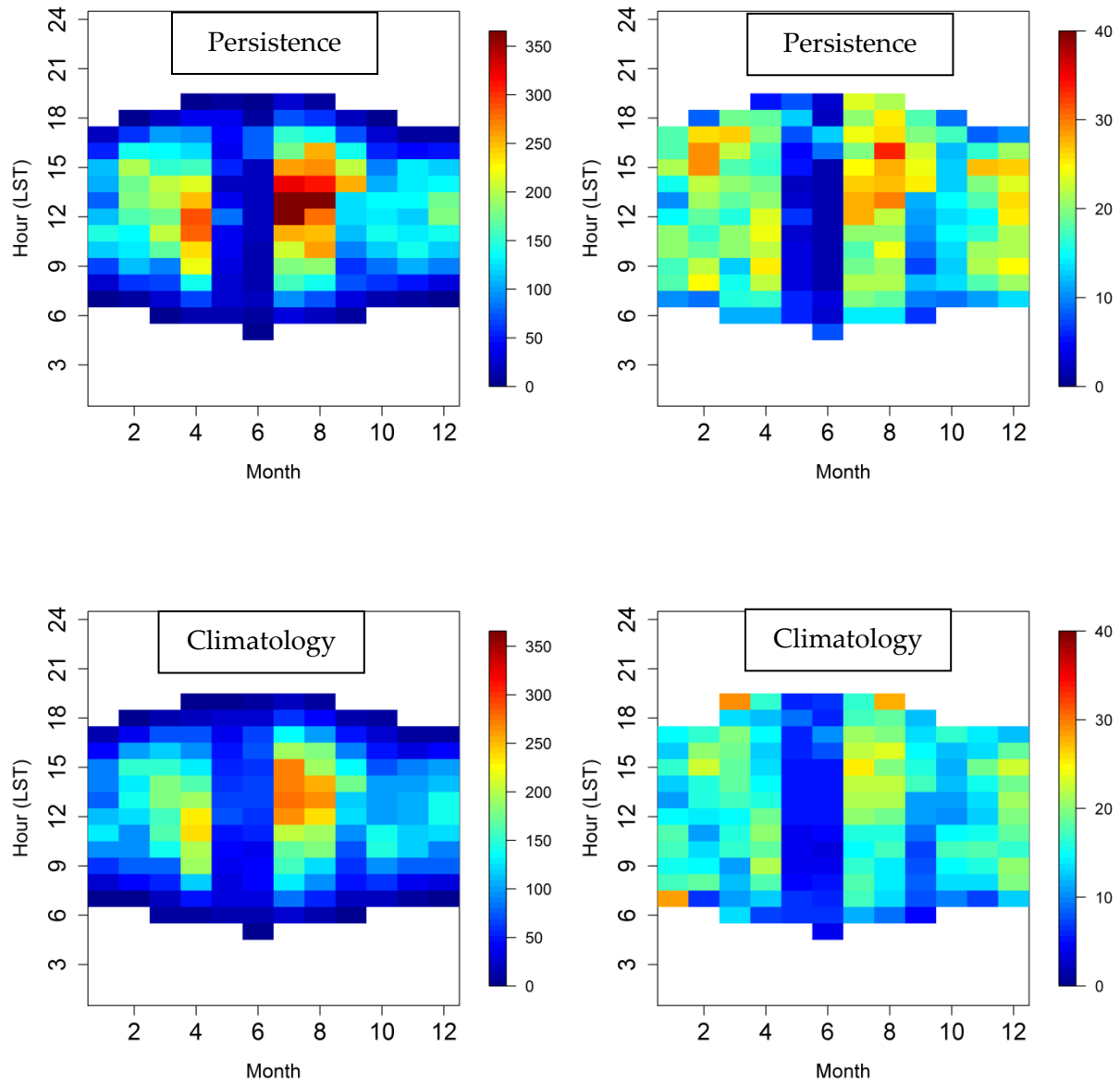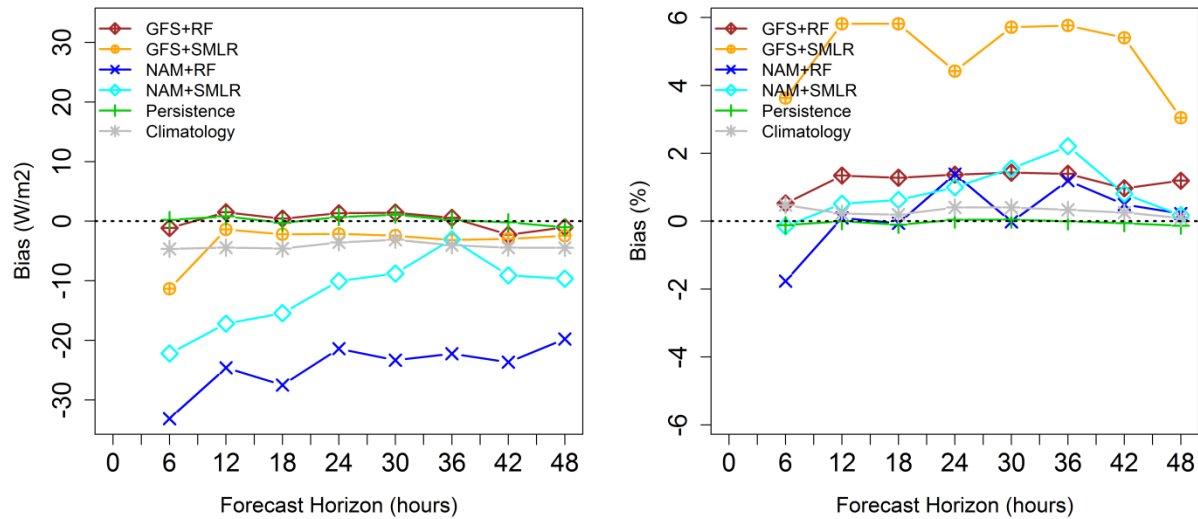

Source: AWS Truepower

All forecast methods were compared over concurrent period of time. Nighttime values are not included.

The RMSE by look-ahead period has similar oscillations as the bias of the hourly-averaged GHI forecasts. Again, the 6-hour block averaging gets rid of the oscillations. Figure 36 shows that the NAM with SMLR performs best and the 24-hour persistence performs the worst. Using climatological GHI forecasts as a reference, all NWP-based forecasts are more accurate and provide value with the exception of the GFS with SMLR. Similar to other research studies (Lara-Fanego et al. 2012, Mathiesen et al. 2011, Perez et al. 2010) the error growth with look-ahead time was not significant over the 6 to 48 hour period. Actually, the RMSE of all NWP-based forecasts remained relatively stable throughout this forecast horizon.

80

Figure 36: RMSE (left panels) and relative RMSE (right panels) by forecast horizon for GHI forecasts averaged over a 6-hour block.
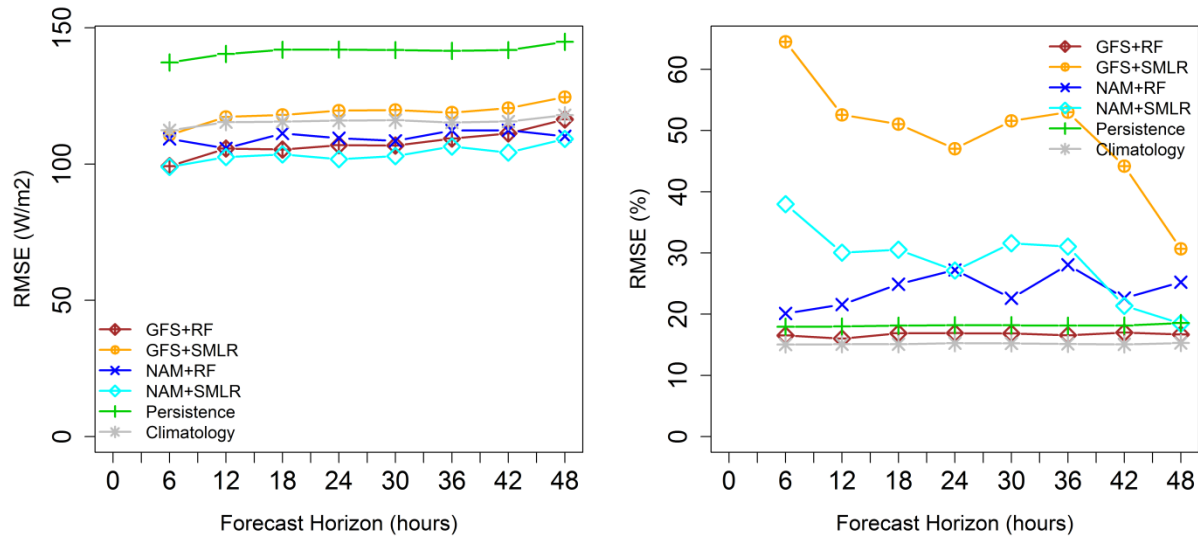
Source: AWS Truepower
All forecast methods were compared over concurrent period of time. Nighttime values are not included.

The relative RMSE plots provide a different view of the performance of each forecast. Out of all the NWP-based forecasts, only the GFS with RF is roughly as good as the climatological forecasts where all others perform worse. The relative mean bias and RMSE statistics give as much weight to transitional hours with low irradiance around sunrise and sunset as midday hours with large irradiance. In practice, forecasting errors around sunrise and sunset are of minimal impact given that the power generation is low when GHI is small.

Taylor diagrams (Taylor 2001) can be very effective at summarizing forecast errors where the best forecasts will lie closest to the observed data point located on the X axis. All the NWP-based and climatological forecasts are clustered closely together in Figure 37. The NAM with SMLR performed the best; however the difference is small when compared to the GFS with RF forecasts.

Figure 37: Taylor diagram of GHI forecasts for the 6 to 48 forecast horizon.
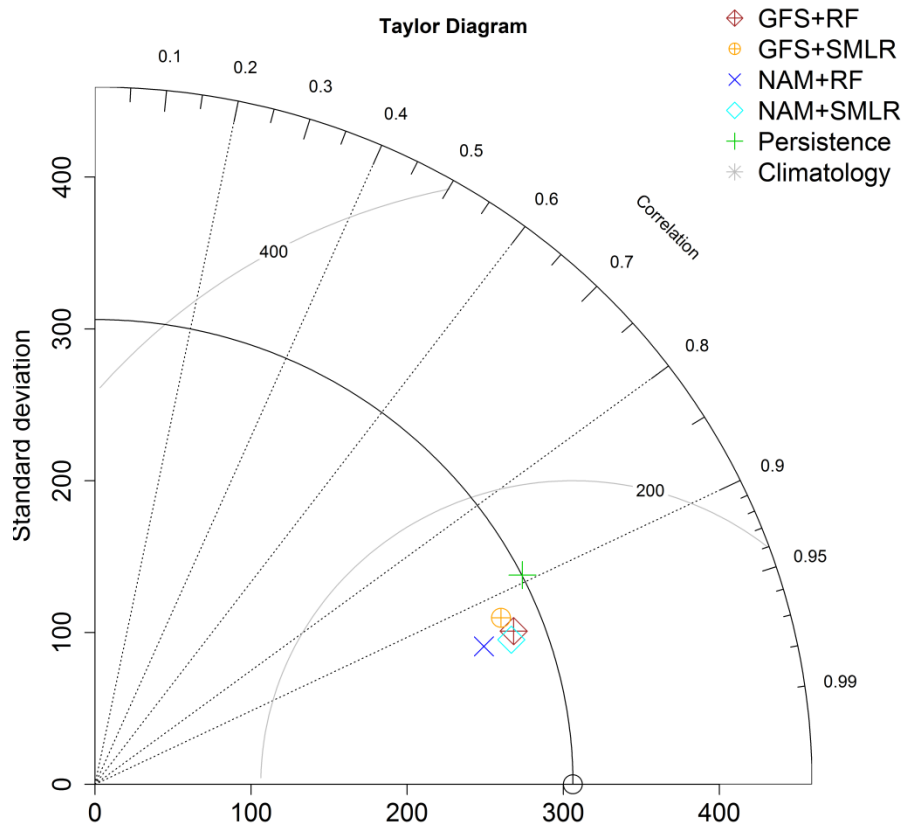
Source: AWS Truepower

The angle corresponds to the correlation coefficient (0 to 1), the radii the standard deviation (W/m2) within each forecast dataset while the light grey arcs refer to the RMSE (W/m2). The best forecasts will lie as close as possible to the observed data point (black circle). All forecast methods were compared over concurrent period of time. Nighttime values are not included.

6.6   Composite Forecast Evaluation

In this section, the performance of the composite forecast system results from integrating four state-of-the-art forecasting methods, namely the rapid update and standard cycle NWP, satellite-cloud motion vector, and TSI-based-cloud motion vector. The composite forecast consists of a blend of different forecasting methods with a heavier weighting given to the forecast method that performed best for each look-ahead period. The composite forecast results relied on an inverse RMSE weighting to blend all the forecasting methods.

Two composite forecasts were developed: (1) a very-short term forecasts (0 to 6 hour lead time) which included the satellite, TSI, persistence as well as the RAP method, and (2) the day-ahead forecasts (6 to 48 hour lead time) composed of the GFS and NAM models. Note that the TSI forecasts were available for a limited number of days (20) in 2012. For 346 days of the hindcast

period of 2012, a persistence forecast based on clearness index was substituted for the missing TSI forecast. The composite forecasts were compared to the individual forecasts in order to confirm the improved accuracy of an ensemble approach to solar irradiance (and power generation) forecasting.

For the hour-ahead time scale (0-6 hour horizon), the composite forecast has lower mean bias and RMSE values on average than any individual methods, although some individual forecasts outperform the composite at specific look-ahead time (Figure 38 and Figure 39). For instance, when looking at the RMSE of the GHI forecasts, the TSI/persistence forecasts do better than the composite at a lead time of 10 minutes by a small margin. The satellite forecasts show the same behavior up to a 2-hour lead time. Overall, the composite forecast performs better than the two reference forecasts, i.e. persistence and climatology. Basically, the composite forecast follows closely the performance of the TSI/persistence forecasts at 10-minute lead time, then the satellite-based forecasts for the 20 minutes to 3-4 hour look-ahead period and to then switch to the NWP RAP method for the remaining time period.

Figure 38: Mean bias (left panels) and relative mean bias (right panels) of GHI forecasts by look-ahead period between 10 minutes to 5 hours.

The composite GHI forecast is shown in yellow. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

Figure 39: RMSE (left panels) and relative RMSE (right panels) of forecasts of the 1-hour average GHI for the 10-minute to 5-hour look-ahead period.

The composite GHI forecast is shown in yellow. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

For the day-ahead time scale (6-48 hour horizon), the composite forecast has a slightly higher mean bias error than some individual methods (see Figure 40) but its RMSE performance is consistently better (see Figure 41, left panel), even compared to the NAM with SMLR, the best performing forecast method tested for this look-ahead period. This is not the case when examining the relative RMSE (see Figure 41, right panel). In fact, the composite forecast is on par with the NAM forecasts but outperformed by GFS with RF as well as the two reference forecasts according to the relative RMSE metric. In any case, the composite forecast for GHI performs better than the reference forecasts, i.e. 24-hour persistence and climatology, and follows closely the best performing individual forecast method for each look-ahead period.

Figure 40: Mean bias (left panels) and relative mean bias (right panels) of forecasts of the 1-hour average GHI by look-ahead period between 6 to 48 hours.
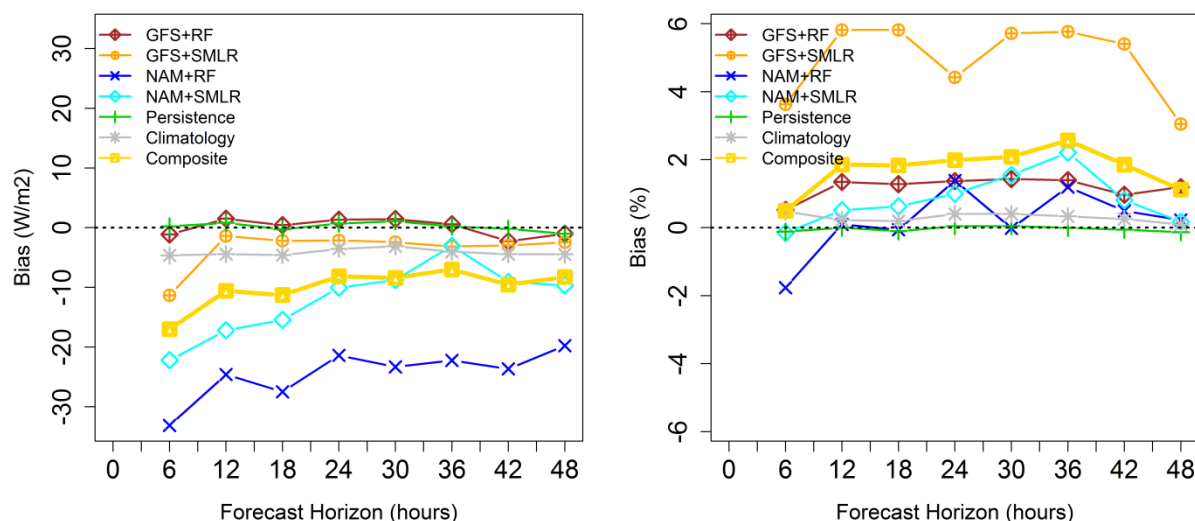
Source: AWS Truepower

The composite GHI forecast is shown in yellow. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

Figure 41: RMSE (left panels) and relative RMSE (right panels) of forecasts of the 1-hour average GHI for the 10-minute to 5-hour look-ahead period.

The composite GHI forecast is shown in yellow. All forecast methods were compared over concurrent period of time. Nighttime values are not included.

Overall, the hour-ahead and day-ahead composite forecasts were able to take advantage of the best performing methods at each look-ahead time scale. We consider this achievement an important step towards the development of an optimal integrated forecast system that can produce a state-of-the-art forecast for all look-ahead time frames of interest. Future research efforts will certainly advance the concept of better performing composite forecasts.

CHAPTER 7:
Conclusions and Recommendations

7.1 Conclusions

The cost and uncertainty associated with the integration of utility-scale solar power can be reduced through advancements in forecasting techniques. The cost-effective management of solar power variability requires high forecast accuracy on a wide range of time scales ranging from minutes ahead to many days ahead. Forecasting methods targeted for specific look-ahead time scales have been developed and are in operational use. Previous research and operational experience has indicated that there is no single forecast method that is optimal for all relevant look-ahead periods (Perez et al. 2010). Although some forecasting entities do rely on more than one method to generate solar forecasts over a broad range of time scales (10 minutes to 48 hours), this study designed and evaluated a composite forecast tool based on an optimization procedure that included several state-of-the-art forecast methods. Four approaches that have previously demonstrated considerable skill for solar forecasting were included in the composite forecast: (1) sky imager - cloud motion vector, (2) satellite - cloud motion vector, (3) rapid update NWP, and (4) standard cycle regional and global NWP. In order to establish a baseline for the validation, reference forecasts were generated using either a persistence or a climatological model based on the observed GHI and actual power production data from the target solar plant.

This study resulted in a number of unique findings that will facilitate advancements in solar forecasting techniques, grid integration and future research opportunities for a composite forecast system design. Major conclusions were:

The intra-hour forecast horizon (10 to 30 minutes) based on 10-minute average forecasts for a 30-day validation period:

- The persistence forecasts achieved the lowest RMSE over the entire period. The persistence forecast assumes that the clearness index, not the GHI or power generation, will remain constant in time. Persistence forecasts are generally used as a reference, but cannot predict ramp events caused by clouds.

- Satellite-based forecasts become as accurate as persistence forecasts at a 30-minute look-ahead interval. Since the satellite forecast error increase slowly with forecast horizon, satellite forecasts will typically outperform persistence over longer look-ahead times.

- The TSI-based forecasts did not outperform the persistence and satellite forecasts. The shadowband on the TSI instrument blocks a significant portion of the sky (~ 14%), which limits its ability to track clouds and advect them in time. Recent advances in sky imaging technologies can reduce the shadowband blocking and are expected to improve the forecast.

The hour-ahead forecast horizon (30 minutes to 5 hours) based on 10-minute average forecasts for the 2012 calendar year:

- The satellite-based forecast outperformed the persistence forecasts. The satellite forecasts achieved the lowest RMSE up to a 3 to 4 hour look-ahead period.

- The rapid update NWP forecasts, i.e. RAP model corrected by a MOS, outperformed the satellite forecasts for the 3.5 to 5 hour look-ahead time.

- Although the satellite-based and rapid update NWP model methods performed better than persistence and climatology for this time scale, the satellite-based forecast errors (RMSE) grew with lead time to reach a similar level of accuracy as the climatology forecasts at a 5-hour look-ahead time.

- Beyond a 2-hour lead time, climatology forecasts produced a lower RMSE than the persistence forecasts.

- All forecasting methods performed significantly better during the months of May and June when the prevailing conditions at the Copper Mountain Solar 1 plant were clear sky. Furthermore, the forecasts performed worse during the months of July and August than any other months throughout the year. This appears to be due to highly variable small-scale cloud regimes associated with the North American monsoon that initiated in July and dissipated in September.

The day-ahead forecast horizon (6 to 48 hours) based on hourly average forecasts for the 2012 calendar year:

- All NWP-based forecasts were more accurate (in terms of RMSE) than the climatological forecasts with the exception of the GFS with SMLR.

- The error (RMSE) growth with look-ahead time remained relatively modest throughout the forecast horizon for all NWP-based methods. This finding was also shown in previous research studies (Perez et al. 2010, Lara-Fanego et al. 2012, Mathiesen and Kleissl 2011).

- The NWP and climatology forecasts performed significantly better during the months of May and June and significantly worse during the months of July and August than any other months throughout the year. Again, this appears to be specific to the target site which was impacted by the North American monsoon initiated in July and dissipated in September.

The hour-ahead (0 to 6 hours) and day-ahead (6 to 48 hours) composite forecast systems based on the 2012 calendar year:

- Both composite forecasts had lower mean bias error and RMSE than any individual methods on average.

- The hour-ahead composite forecasts follow closely the performance of the TSI/persistence forecasts at the 10-minute lead time, satellite-based forecasts for the

20 minutes to 3-4 hour look-ahead period and finally RAP model forecasts for the 3.5 to 5 hour period.

- The day-ahead composite forecasts were well aligned with the best performing NWP model forecasts in this study, i.e. NAM with SMLR.

Results of the composite forecast experiment indicate that constructing a method to capitalize on the best performing forecasting techniques at various look-ahead periods provides a tangible reduction in errors. Therefore, it is likely that future research efforts will benefit from refined component methods and advance the concept of a better performing composite. The results of this project represent an important first step towards the development of an optimal integrated forecast system that can produce a more accurate forecast for all look-ahead time frames of interest than any single forecasting method. While this project was focused on utility-scale solar, most of the techniques are also applicable to distributed solar.

7.2 Recommendations

7.2.1 Universal Dataset

The differences in individual datasets available for each forecast method posed several challenges to the validation of individual forecast results, as well as the integration of these results into a composite forecast. Differences observed included the total plant capacity, data availability and frequency of data output (1-second to 10-minute averages), which complicated the generation of individual forecasts and their integration into the composite forecast system. The common data frequency interval for all forecast methods was 10 minutes, and therefore, this data limitation defined the minimum forecast interval used. The initial phase of future research efforts should focus on assembling a single power production and irradiance dataset from the target generation facility with a format and content from the target plant for research purposes. This would mitigate data limitations that significantly affected the minimum composite forecast interval, composite forecast design, results, and validation of individual methods.

7.2.2 From Hindcast to Real-Time Forecasting

This project was conducted in hindcast mode, using historical data to construct a forecast. Limitations regarding the acquisition of power generation and irradiance data from the target site complicated the design and testing of the composite system and necessitated focus on the development of a hindcast based on the calendar year 2012. However, the generation of a year's worth of power and irradiance data proved challenging as the forecasting techniques are in different states of development. The potential for a real-time deployment of a composite forecast system utilizing all four methods would require a continuous and reliable flow of data from each forecast method as well as a universal power production and irradiance dataset from the target plant itself.

An advantage of real-time forecasting over hindcasting for the NWP-based forecast system is that the real-time NCEP model outputs include GHI whereas this field is missing in the archived datasets for some of the NCEP models, which probably hurts the performance of the NWP forecasts in hindcast mode. However, future hindcast experiments provide tangible

information that can be applied in a composite forecast design as individual methods are further refined.

7.2.3 Next Generation Individual Forecast Method Updates

The results of this project represent an important first step towards the development of an optimally integrated forecast system that can produce a more accurate forecast for all look-ahead time frames of interest, than any single forecasting method. Therefore, it is likely that future research efforts to refine forecast methods will benefit and advance the concept of a better performing composite. Challenges experienced, future research opportunities and planned improvements for the each forecast method are summarized below.

7.2.3.1 Advancements in Sky Imager-Based Forecasts

On-site data collection efforts at the target plant were challenged by data availability for training and testing purposes. Power output and irradiance data (2012) were examined and 60 days with variably cloudy conditions were selected for sky imager model testing. Most non-weather exclusions were attributed to dirt on the imager (22%), indicating more rigorous maintenance is required to achieve high data recovery. However, a significant improvement in forecast quality is expected with new high-dynamic range sky imaging technology that will eliminate the shadowband (Urquhart et al. 2013). The shadowband limited the imager's ability to detect mid to high altitude clouds such as those typically observed at the target site. In addition, future research efforts should exploit opportunities presented by the colocation of two imagers. Two sky imagers not only increase the areal coverage of the images, but also provide opportunities for cloud height determination through stereography. Further research is warranted to determine this method's performance on very short-term time scales.

Although the real-time application of this technique, is challenged by the lack of automated determination of thresholds for cloud detection and currently relies on manual data analysis machine learning tools would likely result in significant improvements to forecast performance and its potential for real-time application.

7.2.3.2 Advancements in Satellite-Based Forecasts

The satellite - cloud motion vector approach provided hindcasts of irradiance for the 2012 calendar year. The current forecasting system can be configured to forecast power generation but this capability was not available at the start of the project. As mentioned earlier, the various forecasting techniques were in different states of development. Recent enhancements to satellite-based methods include the use of multi-channel GOES data to provide pre-dawn satellite-based irradiance forecasts. Additionally, multi-channel GOES data are used to improve forecasts over challenging bright desert and snow covered regions. Future satellite-based forecast advancements will include non-static cloud advection and probabilistic methodologies.

7.2.3.3 Advancements in NWP-Based Forecasts

The performance of forecasts from NWP-based methods is dependent on four primary factors:

(1) The quality of the mathematical representations of atmospherics physics contained in the NWP model itself.

(2) The amount and quality of data from atmospheric sensors that are available to initialize an NWP forecast.

(3) The effectiveness of the NWP data assimilation scheme in blending spatially sparse set of raw atmospheric measurement data from a diverse set of sensors into an accurate representation of the atmospheric state at the time that an NWP forecast is initialized.

(4) The ability of statistical algorithms to reduce the magnitude of systematic errors in the NWP forecasts by operating on the output data from the NWP systems (as procedure typically know as Model Output Statistics or MOS).

Much of the NWP-related effort in this project was directed toward how to most effectively reduce the magnitude of the systematic errors in the NWP forecasts via post processing of the output dataset (#4).The benefit obtained from this post-processing procedure is typically related to two key factors: (1) the size and representativeness of the training sample available to diagnose NWP systematic errors and (2) the effectiveness of the statistical approach in distinguishing among significant NWP error regimes. Several issues were encountered that provide valuable lessons learned and opportunities for future advancements in forecasting methods.

The size and representativeness of the training sample for one of the key NWP models employed in the project was challenged by a discontinuous data set. During the project effort, the RAP model replaced the RUC as NOAA's rapid update model during the 2012 experimental period (on May 1, 2012). This posed a problem for the training of the statistical models (i.e. MOS) since an extended period of forecasts from the new operational NOAA model (i.e. RAP) was not available. For statistical model training purposes of this project, no distinction was made between forecasts generated by the RUC model and those generated by the RAP. However, it is likely that the systematic errors between NOAA models were substantially different. Although this was a one-time event, it could have potentially biased the results. In the future, the hindcast period for training and validation is likely not to be challenged by this limitation.

A second representativeness issue was related to the output variables that were available from each of the models. Some NWP model data available from the government archives did not include forecasted GHI values. Specifically, the forecasted GHI field was available for the NAM model, but not for the GFS, RUC or RAP models. This is a key deficiency for NWP-based forecasts since GHI is generally the most significant predictor when training the statistical models. Therefore, an algorithm was developed to estimate the forecasted GHI for the GFS, RUC and RAP models from the variables that were available in the data archives.

A third challenge was related to NWP performance factors and the amount of data available to initialize forecasts. The RUC and RAP are updated hourly but only every other cycle, i.e. 2-hour cycle, was used during this study to limit the amount of processing and data storage. The use of less frequent updates effectively reduces the amount of observational data that is used in the forecast process. It is recommended that future projects use the highest possible update frequency available from an NWP system.

A fourth issue was related to NWP performance factors #1 and #3. The use of forecasts from a single NWP system results in a high level of performance exposure to deficiencies in a specific model's representation of future conditions. The use of an ensemble of different models reduces exposure to these deficiencies. Regional NWP models as WRF, ARPS and MASS can be configured by forecasting consultants with expertise in numerical modeling. The advantage of running a regional NWP model in-house in addition to using NCEP's models is the ability to fine tune and zoom in on the region of interest with finer resolution and customize the formulation of the model physics for a particular application such as solar irradiance forecasting. In addition, other options are available to obtain forecasts from regional rapid update NWP models. NOAA has been experimenting with the (HRRR), which is based on a configuration of WRF using 3 km grid spacing and updated hourly in real time. The operational implementation of HRRR at NCEP is planned for late spring 2014. Future efforts to advance forecasting methods, especially if performed in real-time, are recommended to include NOAA's latest state-of-the-art NWP forecast system, HRRR. Moreover, AWS Truepower has configured its in-house models to operate in rapid update mode with the assimilation of cloud information derived from satellite imagery and Doppler radar data. The assimilation of this data has the potential to incorporate some of the short-term predictive skill found in satellite methods cloud motion vector model into the NWP methods. Moreover, improvements in forecast error during the monsoon season may be achieved with additional training data, advanced MOS approaches and improved formulations of model physics.

Lastly, the NWP performance could be enhanced by the ability to distinguish different NWP model error regimes. A more effective identification of error regimes enables a more precise correction of systematic errors in the NWP forecasts. Recent advances in artificial intelligence and meteorology have yielded advanced statistical techniques such as artificial neural networks, random forest, support vector machines, boosted regressions and analogue ensemble which have a greater ability to differentiate NWP model error regimes and apply regime-specific adjustments to the NWP forecasts. These approaches provide new opportunities to further develop forecasting tools by applying these techniques to advance NWP models.

7.2.4 Advancements in Composite Forecasts

The results of this project represent an important first step towards the development of an optimal composite forecast system that can produce a more accurate forecast for all look-ahead time frames of interest than any single forecasting method. The design and implementation of the composite forecast system has provided valuable lessons related to the successful blending of different forecast methods.

Two different methods were tested to construct a composite forecast, however a wide variety of statistical tools are available to blend the individual forecasts into a composite forecast. A key attribute to be evaluated by more advanced blending techniques are the performance differences amongst the methods that are linked to differences in weather regimes. An effective exploitation of these differences would enable the weighting of the methods to vary according to their expecive relative performance in the weather regime that is expected to be present in the segment of the forecast period under consideration.

There is also an opportunity to obtain more value from the composite forecast system by constructing a probabilistic forecast. The deterministic forecast approach used in this project (and a favorite of most users because of its simplicity) essentially discards the information about time varying uncertainty that is contained in the ensemble of individual forecasts. Initial steps towards the construction of a composite probabilistic forecast were advanced through the use of quantile regression. This capability should be further developed in future forecasting experiments in order to provide additional value to the grid operators about the time varying uncertainty in the forecasts and facilitate the decision making process at the ISOs.

Another potential opportunity is to formulate metrics that more precisely measure the sensitivity of the grid operator decision-making process to forecast error, and then design the composite forecast method to yield the best performance for those metrics. In this project, the focus was on obtaining the lowest RSME, but this is a very general metric that does not assess the value of the forecast for grid management decision-making very precisely. For example, the grid operator is typically much more sensitive to solar forecast error during critical periods when the solar power production is changing very rapidly (i.e. at times of large upward or downward ramps). There could be considerable value in designing a composite forecast method that is customized for best performance during these periods and not necessarily over all time periods. That is, the weighting of the component methods for this purpose might be quite different than the weighting to achieve best overall performance. Indeed, this concept could be extended to a high degree of sophistication because the sensitivity to forecast error has the potential to be differentiated between up and down ramp events, by time of day, and a number of other factors. Thus, there is a potential for future research projects to refine the methods used to construct a composite forecast to optimize the forecast performance attributes (e.g. ability to predict down ramps at times of peak load) that have the greatest impact on the grid management decision-making process.

Although this project focused on utility-scale solar, most of the techniques are also applicable to distributed or non-utility scale solar. However, it should be noted that the results from this project are only for one site in Nevada. The forecast performance characteristics could be substantially different for locations with different cloud and weather regime climatologies. It will be important to test these approaches for other sites with a range of cloud and weather regime climatologies.

7.3 Public Benefits

The primary result of this project is a validated approach to improve solar irradiance and power generation forecast performance across a broad range of time scales from minutes ahead to 48 hours ahead by constructing an optimal composite of predictions produced by a diverse set of forecast methods that have different horizon performance profiles (i.e. at what look-ahead times they perform best). The cost-effective management of solar power production variability requires high forecast accuracy on time scales ranging from minutes to several days ahead. As solar energy production continues to increase at a rapid pace in California, the development of tools that can predict the solar energy production at multiple look-ahead time scales becomes critical to situational awareness and grid operational planning. Through advancements in

forecasting techniques, the cost and uncertainty associated with the integration of utility-scale solar power can be reduced.

By advancing improvements in solar power production forecast performance there are two primary interrelated economic and social benefits:

- Reduction in the cost of integrating solar power production into an electric system

- Facilitation of higher penetration levels of solar power production on a grid system at a specified level of cost and reliability.

Thus, the benefit to the public are derived from a lower electricity cost for a specified level of solar penetration and reliability, and/or a higher level of solar penetration and its associated reduction in the emission of greenhouse gases.

The complex economic and operational structure of electric systems makes quantifying the benefits of forecasting difficult. However, grid integration studies have been conducted to assess the impact of integrating variable energy sources such as wind and solar energy into specific grid systems. One of these studies, entitled the Western Wind and Solar Integration Study, examined the economic benefit of improved wind and solar forecasts on grid operation costs (Lew et al. 2011) for the grid systems within the jurisdiction of the Western Electric Coordinating Council, which includes the California ISO control area as well as the remaining portions of California. This study indicated that a 10% reduction in the mean absolute error of wind power production forecasts would result in an annual reduction in grid integration costs of $28 Million at a penetration level of 14% and an annual reduction of $52 Million at a penetration level of 24%.

While this calculation was not performed specifically for solar forecasts it is reasonable to expect that similar reductions in the error of solar forecasts would yield a similar magnitude of integration cost reductions. Furthermore, it is likely that this and other studies have underestimated the reduction in costs associated with improved forecasts because they have not considered all the possible ways in which improved forecast information can be used to reduce costs. The focus of these studies has mostly been on the impact of more efficient day-ahead commitment of generating units.

The results of this project have the potential to enable the operators of electric systems in California (i.e. California ISO and other entities) to realize a portion of the benefits of improved forecasting that have been estimated in the previously cited grid integration studies. The anticipated reduction in grid operation costs with the levels of solar penetration expected in California in the near future will be realized through a more cost-effective decision-making on multiple time scales ranging for the minutes to hours ahead decisions for "regulation" and "load following" purposes to the more economical day-ahead commitment of generation resources.

Therefore, improvements in forecast performance associated with the results obtained in this project have the potential, if operationally employed, to reduce the price of electricity produced by a system that has a solar power penetration level that is desired to achieve other long term social, economic or environmental objectives.

CHAPTER 8:
Technology Transfer

8.1 Overview

The purpose of this project was to configure, demonstrate and validate a composite solar forecasting system for California that provides short-term forecasts of solar irradiance and solar-based generation. The primary goal was to design a multiple time scale composite solar forecast system that could assist system operators with timely and accurate foreknowledge of 'intra-hour', 'hour-ahead' and 'day-ahead' solar-based generation so that they can economically accommodate the increasing penetration of this type of generation while balancing other resources and maintaining high levels of grid reliability.

This project incorporated four state-of-the-art forecasting methodologies including sky imagers, satellite – cloud motion as well as rapid update and standard cycle NWP model approaches, customized for California's unique climate and utility operating environment. This innovative approach combined the various forecast methods to produce a composite solar forecast system with an emphasis on the "best of" methods specific to each time scale horizon ('intra-hour', 'hour-ahead' and 'day-ahead'). The composite forecast system look-ahead times range from critical short-term ramp identification periods (10 minutes to 30 minutes ahead, "intra-hour") to the load-following operations of 30 minutes to 5 hours ahead ("hour-ahead") and ultimately, the day-ahead time frame of 5 to 48 hours ("day-ahead") that are important for generation unit commitment.

The purpose of this chapter is to describe the technology transfer activities that served to disseminate the knowledge gained, experimental results and lessons learned to the intended users of project results including industry stakeholders, utility operators, CAISO as well as the national and internal public, private and academic research and forecasting community at large.

8.2 Key Activities

8.2.1 Professional Conferences and Publications

Research to improve renewable energy forecasting methodologies is on-going across a wide variety of public, private and academic institutions. This project uniquely combined the best of several approaches from all three sectors to evaluate the benefits of each and discuss lessons learned during the experiment. On-going research is central to continued improvements in forecasting methodologies and the results of this project were presented at two internationally renowned conferences:

- The 3rd International Workshop on Integration of Solar Power into Power Systems held October 21-24, 2013 in London.

- The American Meteorological Society's Annual Meeting held February 2-6, 2014, in Atlanta, Georgia.

The 3rd Workshop on Integration of Solar Power into Power Systems provided a forum for international recognition of the work being performed in the CEC PIER research program. This event was widely attended by industry experts and scientists studying the challenges associated with grid integration and renewable energy forecasting. The AMS Annual Meeting is the premier event for the presentation of meteorological research amongst atmospheric scientists.

These conferences were used by the project team to introduce the concept of an optimal composite forecast as well as the underlying methodologies and results. Lessons learned and research recommendations were shared with fellow research scientists and their feedbacks helped shape this solar forecasting experiment. Podium presentations were provided and made available to the public. As required by the 3rd International Workshop organizing committee, a paper was also published in the conference proceedings.

8.2.2 Grid Integration and Applicability

AWS Truepower has a unique relationship with CAISO as the wind and solar power forecast provider for the Participating Intermittent Resource Program (PIRP). This service includes hour and day-ahead forecasts for wind and solar plants that participate in PIRP as well as a number of generation facilities designated as Eligible Intermittent Resources.

CAISO, a project partner and key stakeholder, was briefed regularly on the status and challenges of the project to ensure that the forecasting tools and systems refined and integrated during the project had the potential to provide value in the grid management decision-making processes at CAISO. CAISO has provided input on the operational decision-making time scale s and indicated that at this time, operators are limited to decision-making no shorter than 10 minutes ahead. The knowledge gained in this project will be used directly by AWS Truepower to improve the performance of the operational hour-ahead and day-ahead solar power production forecasts provided to CAISO. These include the forecasts that are being produced to address the needs of the special program market transactions as well as the requirements of FERC Order 764.

Many of the component methods used in this project are already being employed in the ensemble forecast system used for the CAISO application. The results obtained from this project will provide the basis for improvements in two areas: (1) refinements to the component methods already being employed in the CAISO forecast production with a focus on an improved MOS procedure to correct system errors in the NWP forecasts; and (2) a more optimal way to create the composite forecast from the individual ensemble members. It is expected that these refinements will result in a substantial reduction in forecast error on both the hours-ahead and day-ahead time scales.

8.2.3 PIER Lessons Learned Information Exchange Workshop

The Energy Commission research and development program has funded many new initiatives with a wide variety of researchers who are seeking to better understand, integrate and manage the generation of renewable energy on the grid system.

The Commission intends to host a workshop with the researchers who completed work under this funding cycle to share results, lessons learned and future research needs with them during the third quarter of 2014. The identification and documentation of the lessons learned are important to the success of future projects, fosters continuous improvement and will serve to help define future research needs and areas of study. AWS Truepower intends to participate in this workshop to share our knowledge of the challenges, significant results and potential for future research opportunities.

GLOSSARY

Term	Definition
ARPS	Advanced Regional Prediction System
CEC	California Energy Commission
CPR	Clean Power Research
CSI	California Solar Initiative
DNI	Direct Normal Irradiance
GHI	Global Horizontal Irradiance
GOES	Geostationary Operational Environmental Satellites
GW	Gigawatt
HRRR	High Resolution Rapid Refresh
k_t	Clearness Index
LST	Local Standard Time
MADIS	Meteorological Assimilation Data Ingest System
MAE	Mean Absolute Error
MASS	Mesoscale Atmospheric Simulation System
MBE	Mean Bias Error
METAR	Meteorological Terminal Aviation Routine Weather Report
MOS	Model Output Statistics
MW	Megawatt
NAM	North American Mesoscale
NAM-LR	NOAA North American Mesoscale model with a Screening Multiple Linear Regression MOS method used in this study
NAM-RF	NOAA North American Mesoscale model with a Random Forest MOS method used in this study
NCEP	National Centers for Environmental Prediction
NDA	Non-Disclosure Agreement
NOAA	National Oceanographic and Atmospheric Administration

NOMADS	National Operational Model Archive & Distribution System
NWP	Numerical Weather Prediction
NWS	National Weather Service
PIER	Public Interst Energy Research
PIRP	Participating Intermittent Resource Program
POE	Probability of Exceedance
PV	Photovoltaic
RAP or RR	Rapid Refresh models
RAP-LR	NOAA Rapid Refresh model with a Screening Multiple Linear Regression MOS method used in this study
RAP-RF	NOAA Rapid Refresh model with a Random Forest MOS method used in this study
RBR	Red-Blue Ratio
RD&D	Research, Development, and Demonstration
RH	Relative Humidity
RF	Random Forest
RMSE	Root Mean Square Error
R2	Coefficient of Determination
Sat-CMV	Cloud motion vector technique used in this study based on satellite visible image data
SMLR	Screening Multiple Linear Regression
RUC	Rapid Update Cycle
TOA	Top of Atmosphere
TSI	Total Sky Imager
TSI-CMV	Cloud motion vector method used in this study based on Total Sky Imager data
UCSD	University of California- San Diego
UTC	Universal Time Coordinate
WRF	Weather Research and Forecast

REFERENCES

Administration National Weather Service, National Centers for Environmental Prediction 4700 Silver Hill Road, Mail Stop 9910 Washington D.C. 20233-9910. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf

Beaucage, P., M.C. Brower, J.D. Frank, J.D. Freedman, 2014: Development of a Stochastic-Kinematic Cloud Model to Generate High-Frequency Solar Irradiance and Power Data. submitted to *Solar Energy* journal.

Boland, J., B. Ridley and B. Brown, 2008: Models of diffuse solar radiation. *Renewable Energy*, vol. 33, 575-584.

Breiman, L., 1984: Classification and Regression Trees, Chapman & Hall, 358 pp.

Breiman, L., 2001: Random Forests. *Machine Learning*, vol. 45, 5-32.

Chow, C., B. Urquhart, A. Dominguez, J. Kleissl, J. Shields and B. Washom, 2011: Intra-Hour Forecasting with a Total Sky Imager at the UC San Diego Solar Energy Testbed. *Solar Energy*, vol. 85, 2881--2893.

DTC, 2013: Gridpoint Statistical Interpolation (GSI): Version 3.2 User's Guide. National Center for Atmospheric Research, National Centers for Environmental Prediction, NOAA Global Systems Division, Earth System Research Laboratory, NOAA [Available from : http://www.dtcenter.org/com-GSI/users/index.php].

Kaplan, M. L., J. W. Zack, V. C. Wong and J. J. Tuccillo, 1982: Initial results from a mesoscale atmospheric simulation system and comparisons with an AVE-SESAME I data set, *Mon. Wea. Rev.*, vol. 110, 1564-1590.

Glahn, H. R., and D. A. Lowry, 1972: The use of ModelOutput Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1202-1211.

Gao, J., M. Xue, K. Brewster, F. Carr and K. K. Droegemeier, 2001: A Three-Dimensional Variational Data Assimilation Scheme for storm-scale model, Preprints, 14th Conference on Numerical Weather Prediction, 30 July – 2 August 2001, Amer. Meteorol. Society. pp. 72-74.

Gao, J., M. Xue, K. Brewster, F. Carr and K. K. Droegemeier, 2002: New developments of a 3DVAR system for a nonhydrostatic NWP model, Preprints, 15th Conf. on Numerical Weather Prediction, 12-16 August, Amer. Meteorol. Society. pp 339-342.

Gao, J., M. Xue, K. Brewster and K. K. Droegemeier, 2004: AThree-dimensionalVariational Data Assimilation Method with recursive filter for Single-Doppler Radar, *J. Atmos. Oceanic. Technol.* vol. 21, 457-469.

Global Climate and Weather Modeling Branch, 2003: The GFS Atmospheric Model. NCEP Office Note, No. 442. U.S. Department of Commerce, National Oceanic and Atmospheric Administration National Weather Service, National Centers for Environmental

Prediction 4700 Silver Hill Road, Mail Stop 9910 Washington D.C. 20233-9910 [Available from http://www.emc.ncep.noaa.gov/officenotes/newernotes/on442.pdf]

Hammer, A., D. Heinemann, C. Hoyer and E. Lorenz, 2001: Satellite Based Short-Term Forecasting of Solar Irradiance - Comparison of Methods and Error Analysis, Proc. EUMETSAT Meteorological Satellite Data Users Conf.

Ineichen, P. and R. Perez, 2002: A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*, vol. 73, 151–157.

Janjic, Z., R. Gall, and M.E. Pyle, 2010: Scientific Documentation for the NMM Solver, NCAR Technical Note NCAR/TN-477+STR. http://nldr.library.ucar.edu/collections/technotes/asset-000-000-000-845.pdf

Jurado, M., J. Caridad and J. Ruiz, 1995: Statistical distribution of the clearness index with radiation data integrated over 5 min intervals, *Solar Energy*, vol. 55, 469-473.

Lara-Fanego, V., J.A. Ruiz-Arias, D. Pozo-Vázquez, F.J. Santos-Alamillos, J. Tovar-Pescador, 2012: Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain), *Solar Energy*, vol. 86, 2200-2217.

Lew, D., M. Milligan, G. Jordan and R. Piwko. 2011. The Value of Wind Power Forecasting. *Proc. 91st American Meteorological Society Annual Meeting, the Second Conference on Weather, Climate and the New Energy Economy. Washington, D.C.*

Lorenz, E., D. Heinemann, H. Wickramarathne, H.G. Beyer and S. Bofinger, 2007: Forecast of Ensemble Power production by Grid-connected PV Systems. *Proc. 20th European PV Conference, Milano, Italy*.

Mass, C.F., D. Ovens, K. Westrick, and B.A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bulletin of the American Meteorological Society*, 407-430.

Mathiesen, P.,J. Kleissl , 2011: Evaluation of numerical weather prediction for intra-day hourly solar irradiance forecasting in the CONUS, *Solar Energy*, vol. 85, 967-977.

Page, J., 2003: The role of solar radiation climatology in the design of photovoltaic systems. Practical handbook of photovoltaics: fundamentals and applications, T. Markvart, L. Castaner, Editors, Elsevier. pp. 5–66.

Perez, R., P. Ineichen, K. Moore, M. Kmiecikm, C. Chain, R. George and E. Vignola, 2002: "A New Operational Satellite-to-Irradiance Model", *Solar Energy*, vol. 73, 307-317.

Perez, R., S. Kivalov, A. Zelenka, J. Schlemmer and K. Hemker Jr., 2010: Improving the Performance of Satellite-to-Irradiance Models using the Satellite's Infrared Sensors. Proc., ASES Annual Conference, Phoenix, Arizona.

Perez, R., S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renné, and T.E. Hoff, 2010. "Validation of short and medium term operational solar radiation forecasts in the US". Solar Energy, vol. 84, p. 2161-2172.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers, 2005. A description of the Advanced Research WRF Version 2. NCAR Technical Note-468+STR. http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf

Stauffer D. R. and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, vol. 118, 1250–1277.

Stauffer, D.R. and N. L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.*, vol. 33, 416-434.

Taylor, K.E., 2001: Summarizing multiple aspect of model performance in a single diagram, *Journal of Geophysical Research*, vol.106, 7183-7192.

Urquhart, B., M. Ghonima, D. Nguyen, B. Kurtz, C. Chow and J. Kleissl, 2013: Sky Imaging Systems for Short-term Forecasting, in: Solar Resource Assessment and Forecasting (Editor, Jan Kleissl), Elsevier.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, and K.Brewster, 1995: ARPS Version 4.0 User's Guide. Center for Analysis and Prediction of Storms, [Available from CAPS, Univ. of Oklahoma, 100 E. Boyd St., Norman OK 73019], 380 pp.

Xue, M., K. K. Droegemeier and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) – A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Physics.* vol. 75, 161-193.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) – A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Physics.* vol. 76, 143-165.

Yang, H., B. Kurtz, A. Nguyen, B. Urquhart, C.W. Chow, M. Ghonima, J. Kleissl, 2014: Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego, *Solar Energy*, in press.